



HAL
open science

Pauvreté, risques et inégalités : trois essais inspirés de l'économie Haïtienne

Jean-Baptiste Anténord

► **To cite this version:**

Jean-Baptiste Anténord. Pauvreté, risques et inégalités : trois essais inspirés de l'économie Haïtienne. Economies et finances. Université de Lille; Université Quisqueya (Port-au-Prince, Haïti), 2022. Français. NNT : 2022ULILA007 . tel-04724056

HAL Id: tel-04724056

<https://theses.hal.science/tel-04724056v1>

Submitted on 7 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ecole Doctorale Société et
Environnement (EDSE)



Ecole Doctorale Sciences Economiques,
Sociales, de l'Aménagement et du
Management (SESAM, ED N° 73)



Pauvreté, risques et inégalités : trois essais inspirés de l'économie haïtienne

Thèse en co-tutelle présentée par

Jean-Baptiste ANTENORD

Pour l'obtention du grade de Docteur en Sciences Économiques

Membres du Jury

Sophie DABO-NIANG, Professeure, Université de
Lille, **Présidente du Jury**

Etienne BILLETTE de VILLEMEUR, Professeur,
Université de Lille, **Directeur de Thèse**

Raulin Lincifort CADET, PhD, Enseignant-chercheur,
Université Quisqueya, **Directeur de thèse**

Béatrice REY-FOURNIER, Professeure, Université
Lyon 2, **Rapporteuse**

Emmanuel FLACHAIRE, Professeur, Université Aix-
Marseille, **Rapporteur**

Jean Marie CAYEMITTE, PhD, Enseignant-chercheur,
Université d'Etat d'Haïti, **Examineur**

Achis CHERY, PhD, Enseignant-
chercheur, Université d'Etat d'Haïti,
Examineur

Benedique Paul, PhD, Enseignant-
chercheur, Université Quisqueya,
Examineur

Danilia Moïse ALTIDOR, Ancienne
Directrice Générale de l'Institut
Haïtien de Statistique et
d'Informatique (IHSI), **Invitée
d'honneur**

Evens EMMANUEL, HDR, Professeur,
Vice-recteur à la recherche,
Université Quisqueya, **Invité
d'honneur**

Thèse soutenue publiquement le : 5 juillet 2022

L'Université Quisqueya et l'Université de Lille n'entendent donner aucune approbation ni improbation aux opinions émises dans cette thèse. L'auteur est le seul responsable de ces opinions.

Résumé

Résumé. La thèse est composée de trois essais. Le premier est consacré au développement de méthodes statistiques adaptées à des situations où on ne peut pas supposer a priori que l'on dispose d'un échantillon représentatif, mais où on dispose par contre de plusieurs sources de données relatives à cette même population. La méthode dite de "capture-recapture", introduite initialement pour faire du dénombrement, est reprise et étendue pour proposer une procédure de "régression" (économétrique) qui combine les observations de plusieurs échantillons indépendants pour corriger les possibles biais induits par leur non-représentativité. Dans le cadre d'un modèle reliant deux variables binaires, on construit les estimateurs statistiques des probabilités conditionnelles et leur variance est calculée. Ces estimateurs sont asymptotiquement sans biais mais leur variance est supérieure à celle d'un estimateur « naïf », c'est à dire qui ne prend pas en compte les possibles biais de sélection. Un test statistique est proposé pour décider, sur la base de l'importance estimée du biais d'estimation, de l'opportunité d'utiliser ou non notre méthode de redressement pour estimer ce modèle binaire.

Le second essai propose un modèle théorique afin d'élaborer des stratégies de minimisation des pertes résultant des aléas climatiques, pour les petites exploitations agricoles. Plus précisément, il s'agit d'évaluer l'opportunité de recourir à un arrosage manuel, à la survenue d'un épisode de sécheresse, sachant que la récolte n'est jamais à l'abri d'un épisode d'inondation destructeur. On établit que l'arrosage manuel est rentable et augmente le revenu espéré si : a) l'intensité des sécheresses est relativement élevée mais le risque d'inondation est peu important ; b) la corrélation entre risque de sécheresse et risque d'inondation est faible ou même négative; c) les éventuelles compensations mises en place par l'autorité publique en cas d'inondation tiennent compte de l'effort spécifique des exploitants. La garantie d'une aide financière (exogène) et une couverture d'assurance endogène en cas d'inondation ont aussi toutes deux un impact positif sur l'effort d'arrosage. Deux résultats importants sont établis : 1) Il n'est pas possible de trouver un niveau d'effort θ ; et une police d'assurance inondation A tels que le couple (θ, A) qui maximise l'utilité espérée de l'exploitant soit associé à une situation sans risque pour l'exploitant ; 2). Dans certaines situations, il est impossible d'avoir une assurance inondation rentable

si l'exploitant ne fait jamais l'effort d'arroser - et inversement, il n'est jamais rentable de faire l'effort d'arroser, si on ne dispose pas d'assurance inondation.

Le troisième essai s'intéresse à l'impact des inégalités de revenu, sur les marges pratiquées par les entreprises, et donc le niveau des prix – et in fine l'accès aux biens et services pour les ménages les plus pauvres. On propose d'abord un modèle de demande unitaire, que l'on étend légèrement pour incorporer des effets revenus. Dans ce modèle simple, quand la distribution de revenu est une distribution de Pareto, la marge réalisée par un monopole peut être calculée par une formule explicite. Cette marge augmente avec les inégalités et tend au contraire vers zéro quand on s'approche de la distribution égalitaire. Nous offrons ensuite un résultat technique permettant d'établir la Schur-concavité d'une fonction implicite. Dans un modèle plus général, où les individus ont tous les mêmes préférences et diffèrent seulement par leur revenu, nous montrons alors que la marge relative réalisée à l'équilibre par un monopole qui maximise ses profits constitue un index de mesure des inégalités de revenus de cette population. Il existe donc des circonstances assez larges où les inégalités sont source d'inefficacités sur les marchés. Dans ce cas, toute redistribution qui amène une baisse des inégalités est génératrice d'efficacité – indépendamment de toute considération normative.

Abstract. The thesis consists of three essays. The first is devoted to the development of statistical methods adapted to situations where one cannot assume a priori to have a representative sample, but where there are several sources of data related to the same population. The so-called "capture-recapture" method, initially introduced for enumeration, is taken up and extended to propose a (econometric) "regression" procedure that combines the observations of several independent samples to correct the possible biases induced by their non-representativity. In the framework of a model linking two binary variables, statistical estimators of the conditional probabilities are constructed and their variance is calculated. These estimators are asymptotically unbiased but their variance is greater than that of a "naive" estimator, i.e. one that does not take into account possible selection biases. A statistical test is offered to decide, on the basis of the estimated importance of the estimation bias, whether or not to use our adjustment method to estimate

this binary model.

The second essay proposes a theoretical model to develop strategies for minimizing losses resulting from climatic hazards for small farms. Specifically, it assesses the appropriateness of manual watering during a drought event, given that the crop is never safe from a destructive flood. It is established that manual watering is profitable and increases the expected income if: a) the intensity of droughts is relatively high but the risk of flooding is low; b) the correlation between the risk of drought and the risk of flooding is low or even negative; c) the possible compensation provided by the public authority in case of flooding takes into account the specific effort of the farmers. The guarantee of (exogenous) financial support and endogenous insurance coverage in case of flooding also have a positive impact on watering effort. Two important results are established: 1) It is not possible to find a level of effort θ ; and a flood insurance policy A such that the pair (θ, A) that maximizes the expected utility of the farmer is associated with a risk-free situation for the farmer; 2). In some situations, it is impossible to have profitable flood insurance if the operator never makes the effort to water - and conversely, it is never profitable to make the effort to water, if one does not have flood insurance.

The third essay looks at the impact of income inequalities on the margins charged by firms, and therefore on the price level - and ultimately on access to goods and services for the poorest households. We propose first a unitary demand model, which we extend slightly to incorporate income effects. In this simple model, when the income distribution is a Pareto distribution, the margin adopted by a monopoly profit maximizing can be calculated by an explicit formula. This margin increases with inequality and tends towards zero as one approaches the egalitarian distribution. We then provide a technical result to establish the Schur-concavity of an implicit function. In a second and more general model, where all individuals have the same preferences and differ only in their income, we then show that the relative margin realised in equilibrium by a profit-maximizing monopoly is a measure of income inequality in this population. There are thus broad circumstances in which inequality is a source of inefficiency in markets. When this is the case, any redistribution that leads to a decrease in inequality is efficiency-enhancing - regardless of

any normative considerations.

Rezime. Tèz la fèt ak twa esèy. Premye a konsakre l nan devlopman yon metòd estatistik adapte nan sitiyasyon kote yo pa ka sipoze a priori ke gen yon echantiyon reprezantatif ki disponib, men kote gen plizyè sous done ki gen rapò ak popilasyon sa a menm. Nou itilize metòd ki rele "capture-recapture" la, ki te itilize okòmansman pou konte kantite manb ki genyen nan yon popilasyon, pou pwolonje l nan pwopoze yon pwosedi "regresyon" (ekonometrik) ki konbine obsèvasyon plizyè echantiyon endepandan pou korije byè ki soti nan echantiyon yo ki pa reprezantatif. Nan kad yon modèl ki konekte de varyab binè, nou konstwi estimatè estatistik pwobabilite kondisyonèl yo epi nou kalkile varyans yo. Estimatè sa yo, pou anpil anpil done, san byè men varyans yo pi gran pase varyans estimatè "nayif" yo, sa vle di estimatè ki pa pran an kont pwoblèm byè seleksyon posib. Nou pwopoze yon tès estatistik pou deside, ki baze sou enpotans byè estimasyon an, si wi ou non pou nou itilize metòd ajisteman nou an pou estime modèl binè sa a.

Dezyèm esèy la pwopoze yon modèl teyorik pou devlope estrateji pou minimize pèt ki soti nan malè pandye klimatik, pou moun kap travay ti moso tè. Plis presizeman, nou chèche evalye ki benefis ki genyen si ti agrikiltè sa yo fè awozaj manyèl, lè yo sispèk yon sechrès ka rive brit sou kou, san pa bliye tou ke rekòt la pa janm an sekirite paske yon inondasyon ka debake epi detwi tout rekòt la. Nou etabli ke awozaj manyèl la pwofitab epi ogmante revni ti agrikiltè yo espere si: a) entansite sechrès la relativman wo men risk pou inondasyon yo ba; b) korelasyon ki genyen ant risk sechrès ak risk inondasyon fèb ki ka menm negatif; c) nenpòt konpansasyon otorite piblik la mete an plas an ka ta gen inondasyon ki pran an konsiderasyon efò espesifik operatè yo. Garanti asistans finansye (ekzojèn) ak kouvèti asirans andojèn nan lè ta gen inondasyon, tou de gen yon enpak pozitif sou efò irigasyon an. Nou etabli de rezilta enpòtan: 1) Li pa posib pou jwenn yon nivo efò θ ak yon kontra asirans inondasyon A tankou koup (θ, A) sa ki maksimize itilite espere operatè a asosye ak yon sitiyasyon san risk pou operatè a; 2). Nan kèk sitiyasyon, li enposib pou gen yon asirans inondasyon pwofitab si operatè a pa janm fè efò pou wouze - Yon lòt bò, li pa janm pwofitab pou fè efò pou wouze, si pa gen okenn asirans inondasyon.

Twazyèm esèy la konsantre sou enpak inegalite revni sou maj benefis antrepriz yo,

kidonk sou nivo pri yo - finalman sou aksè ak byen ak sèvis pou menaj ki pi pòv yo. Nou pwopoze yon modèl demann pou yon sèl inite, ke nou pwolonje yon ti kras nan enkòpore efè revni. Nan modèl senp sa a, lè distribisyon revni an se yon distribisyon Pareto, nou ka kalkile maj ke yon monopol kap maksimize pwofi l reyalize ak yon fòm eksplisit. Maj sa ogmante ak inegalite epi, okontrè, gen tandans ale nan direksyon zewo lè nou apwoche distribisyon egalite a. Nan menm kad sa a, nou ofri yon rezilta teknik ki pèmèt nou etabli "Schur-konkavite" nan yon fonksyon enplisit. Nan yon modèl ki pi jeneral, kote moun yo tout gen menm preferans epi diferan sèlman pa revni yo, nou montre lè sa a, maj relatif ke yon monopol ki maksimize pwofi li reyalize lè gen ekilib konstitye yon endèks pou mezire inegalite revni nan popilasyon sa a. Se poutèt sa, gen anpil sikonstans kote inegalite se yon sous inefikasite nan mache yo. Nan ka sa a, nenpòt redistribisyon ki mennen nan yon rediksyon nan inegalite yo jenere efikasite - pou nenpòt ki konsiderasyon ki gen rapò ak nòm nan yon sosyete.

Mots clés: Estimateurs capture-recapture, régression binaire, risques climatiques, arrosage à petite échelle et assurance agricole, modèle de demande unitaire avec effets revenus, distribution de revenus et inclusion des plus pauvres.

Remerciements

Je voudrais remercier tout d'abord mes directeurs de thèse les professeurs Etienne Billette de VILLEMEUR et Raulin L. CADET pour m'avoir fait confiance et m'avoir accompagné durant ce parcours vers le doctorat. Leur direction et leur encadrement m'ont été lumineux. Qu'ils trouvent ici l'expression de mes sincères remerciements.

J'adresse des remerciements spéciaux au professeur et directeur de recherche CNRS David CRAINICH qui a su m'assister dans l'élaboration du modèle d'assurance pour les petites exploitations agricoles. Merci professeur David CRAINICH. Mes remerciements vont aussi aux professeurs de l'Université de Lille, Nicolas DEBARSY, Jean-Baptiste DESQUILBET, Stéphane VIGEANT qui m'ont conseillé sur des points spécifiques lors de la recherche doctorale. Mes plus cordiaux remerciements aux doctorant(e)s de la Ruche du LEM pour leur support et leur convivialité.

Que mes anciens professeurs Danilia M. ALTIDOR et Lesly Emmanuel DEJEAN trouvent ici l'expression de ma reconnaissance pour leur support inconditionnel depuis le début de cette aventure.

Je n'aurais pas pu arriver au bout de ce parcours sans l'accompagnement et le support soutenus de l'Université Quisqueya et de mon institution de rattachement le Centre de Techniques de Planification et d'Economie Appliquée (CTPEA) à travers son Directeur Général M. Hosval TRISTANT. Merci au *CTPEA* et à *l'UniQ* d'avoir fait ce choix responsable et exemplaire d'investir dans l'humain et pas "nan limen kawoutchou".

Un grand merci à l'Agence Universitaire de la Francophonie (AUF) pour m'avoir en partie supporté financièrement au cours de mes trois premières années de thèse.

Des remerciements spéciaux à mes collègues Schmied ST-FLEUR et Oscar WALGUEN pour m'avoir suppléé respectivement dans le cadre de mes enseignements au CTPEA et à l'Ecole Normale Supérieure (ENS) lors de mes mobilités en France.

Un merci spécial à ma collègue Carline Joseph Duval pour tout son support et les conseils qu'elle m'a prodigués tout au long de cette aventure.

Mes remerciements vont également à l'ingénieur Ange JEANTY et son épouse Marie Clauda D. JEANTY pour leur support moral et le support qu'ils ont donné à mon épouse et mes deux filles lors de mes mobilités en France.

Un merci de tout coeur à mon épouse Medeline P. J. ANTENORD et à mes deux filles Gypsie La Rossa et Jessika Medgina pour leur affection, leur encouragement, leur accompagnement et leur support constant et inconditionnel.

Enfin j'adresse un grand merci à mes frères et soeurs Elima Jacques ANTENOR, Marie Suzie ANTENOR, Rosena ANTENOR, ma nièce-amie Natacha MOISE, ma nièce Wilnie Moise Doll et mon neveu le Rév. Wilkenson Moise pour leur accompagnement et leur encouragement.

Enfin un grand merci à toutes et à tous qui, d'une façon ou d'une autre, de près ou de loin, ont contribué à la réussite de ce projet de thèse.

Table des matières

1	Régression binaire par la méthode de capture-recapture	8
1.1	Introduction	12
1.2	Méthodes capture-recapture (CR)	14
1.2.1	Définition, but, origines, méthodes et modèles	14
1.2.1.1	Définition et but	14
1.2.1.2	Les origines de ces méthodes	14
1.2.1.3	Méthodes et modèles	15
1.2.2	Approche multiliste	18
1.3	Les principaux résultats du modèle	21
1.4	Le modèle de Régression Binaire par la Méthode « Capture-Récapture »	22
1.4.1	Notations générales et définition	24
1.4.2	Cas de deux listes	27
1.4.2.1	Expression matricielle des données de capture recapture pour deux sources	27
1.4.2.2	Les principaux estimateurs de la taille N dans le cas de deux listes	28
1.4.2.3	Les estimateurs des paramètres du modèle	35
1.4.2.4	Le modèle estimé	45
1.5	Les résultats du modèle eu égard à la littérature	47
1.6	Synthèse des principaux résultats et conclusion	48
	Appendices	54

1.A	Annexes	54
1.A.1	L'expression de l'estimateur \widehat{n}^{yx}	54
1.A.2	Variance asymptotique de \widehat{n}^{yx} établie par Sekar et Deming	54
1.A.3	Cas d'une population à deux états	55
1.A.4	Variance asymptotique de \widehat{n}^{yx}	57
1.A.5	Biais asymptotique de l'estimateur naïf de \widetilde{q}	61
1.A.6	Espérance et Variance de o^{yx}	63
1.A.6.1	Espérance mathématique de o^{yx}	63
1.A.6.2	Variance de o^{yx}	65
1.A.6.3	Comparaison de $V(\widehat{q})$ et de $V(\widetilde{q})$	67
1.A.7	Espérance et Variance de \widehat{n}_{00}^{yx}	69
1.A.7.1	Espérance mathématique et variance de \widehat{n}_{00}^{yx}	69
1.A.7.2	Expression de \widehat{p} et sa variance	71
1.A.7.3	Variance et covariance des estimateurs $\widehat{\beta}_0$ et $\widehat{\beta}_1$	74
2	Arrosage privé à petite échelle : un modèle d'assurance pour les petites exploitations agricoles	77
2.1	Introduction	80
2.2	Les principaux résultats du modèle	82
2.3	Le modèle	83
2.3.1	Corrélation entre les risques de sécheresse et d'inondation et effort d'arrosage	90
2.3.1.1	Risque Relatif, effort d'arrosage et sévérité de la sécheresse	94
2.3.1.2	Impact des revenus sur les efforts d'arrosage	98
2.3.2	Revenu espéré avec et sans arrosage	100
2.3.3	Effort d'arrosage en présence d'assurance sur les inondations.	102
2.3.3.1	Impact d'une assurance exogène sur les inondations	103
2.3.3.2	Couverture d'assurance endogène	107
2.3.3.3	Détermination simultanée d'une couverture d'assurance A et d'un niveau d'effort d'arrosage θ optimaux	114

2.4	Illustrations des principaux résultats	115
2.4.1	Sécheresse d'intensité maximale, sécheresse/inondation et effort d'arrosage	115
2.4.2	Revenu, revenu espéré, sécheresse et effort d'arrosage	117
2.4.3	Couverture d'assurance sur les inondations	121
2.4.3.1	Couverture d'assurance exogène	121
2.4.3.2	Couverture d'assurance endogène	122
2.5	Les résultats du modèle eu égard à la littérature	124
2.6	Synthèse des résultats et conclusion	127
Appendices		135
2.A	Annexes	135
2.A.1	Résultats théoriques	135
2.A.1.1	Impact formel de π_{SI} sur l'effort d'arrosage optimal	135
2.A.1.2	Impact du revenu sur l'effort d'arrosage	136
2.A.1.3	Impact d'une assurance exogène sur les inondations	137
2.A.1.4	Impact de l'effort d'arrosage sur la couverture d'assurance endogène	140
2.A.1.5	<i>Impossibilité</i> d'avoir une assurance inondation rentable en situation de zéro <i>effort d'arrosage</i>	147
2.A.2	Résultats des illustrations	149
2.A.2.1	La paire effort-assurance (θ^*, A^*) optimale et courbes de niveau (q, p)	149
2.A.2.2	Effort d'arrosage θ^* et niveau d'assurance A^* qui maximisent l'utilité espérée de l'exploitant	151
3	Accès aux biens et services des ménages aux revenus les plus faibles : un modèle microéconomique de demande	154
3.1	Introduction	158
3.2	Les principaux résultats du modèle	161

3.3	Le Cadre général	162
3.4	Description du modèle	163
3.4.1	Préférences des ménages et principales hypothèses du modèle	163
3.4.2	Utilité et demande	167
3.4.2.1	Utilité et demande individuelle	167
3.4.2.2	Demande agrégée	168
3.4.3	Profits et Bien-être social	169
3.4.4	Maximisation du profit et du bien-être	171
3.4.4.1	Maximisation du profit	171
3.4.4.2	Maximisation du bien-être	182
3.5	Exclusion	185
3.5.1	Maximisation du profit et exclusion	187
3.5.2	Maximisation du bien-être social et exclusion	188
3.5.3	Illustrations des résultats	188
3.5.3.1	Le cas de la distribution de Pareto	189
3.5.3.2	Distribution log-normale	197
3.5.4	Généralisation des résultats du modèle	202
3.5.4.1	Demande unitaire et revenu	202
3.5.4.2	Caractérisation d'une fonction $R_0(p)$, d'une fonction $D(p)$ et conditions d'accès au bien à tous	203
3.6	Les résultats du modèle eu égard à la littérature	209
3.7	Synthèse des principaux résultats	211
Appendices		219
3.A	Annexes	219
3.A.1	Prix optimal qui maximise le profit	219
3.A.2	Conditions Revenu minimum du ménage préférant la qualité uni- taire X_j à la qualité unitaire X_i	220
3.A.3	Prix optimal qui maximise le bien-être social	222

3.A.4 Caractérisation d'une fonction $R_0(p)$, d'une fonction $D(p)$ et éta-
blissement d'une condition de non-exclusion 223

Table des figures

2.4.1 Effort d'arrosage, aversion au risque et sécheresse maximale	116
2.4.2 Comportement de l'effort d'arrosage θ , du coefficient d'aversion au risque η suite à la variation de p et π_{SI}	117
2.4.3 Revenu espéré, Probabilité sécheresse et inondation et effort d'arrosage . .	119
2.4.4 Revenu espéré, probabilité sécheresse et inondation et effort d'arrosage . .	120
2.4.5 Revenu espéré, probabilités sécheresse et inondation et effort d'arrosage .	121
2.4.6 Probabilité de sécheresse et probabilité conjointe sécheresse/inondation . . .	122
2.4.7 Probabilités de sécheresse et d'inondation, assurance endogène et effort d'arrosage optimaux : $\pi_{SI} \in \{1/8, 1/4, 3/8, 1/2\}$	123
2.4.8 Probabilités de sécheresse et d'inondation, assurance endogène et effort d'arrosage optimaux : $\rho \in \{1/4, 5/12, 7/12, 3/4\}$	124
3.5.1 Prix p^* et marge m^* du monopole en fonction de l'indice de Gini \mathcal{G}	191
3.5.2 Paramètre α , Prix optimal p^* et Revenu R_0	196
3.5.3 $\bar{E}(\underline{R})$ et $D(\underline{p})$ en fonction de \underline{R}	197
3.5.4 Préférences $v(\cdot)$ des ménages	204
3.5.5 Comportement de $h(R)$ et $w(R)$	208

Liste des tableaux

1.1	Historique de capture de N individus sur K listes	19
1.2	Matrice des résultats de présence/absence pour deux listes	28
2.3.1	Les différents états de la nature et les probabilités associées	84

Introduction Générale

Le caractère multidimensionnel de la pauvreté est universellement reconnu par les principaux chercheurs travaillant sur la question. De sorte que l'essentiel des récentes élaborations théoriques et conceptuelles relatives à la pauvreté et les politiques de lutte contre celle-ci s'effectue autour de ce caractère multidimensionnel. Celui-ci se retrouve déjà dans les travaux de Rawls (Dieterlen, 2005) qui conceptualise la pauvreté en tant que carence en biens primaires¹. Selon le *PNUD*, la pauvreté est un phénomène complexe et multidimensionnel (Benicourt, 2001). Cette conception du *PNUD* de la pauvreté ne se diffère pas trop de celle de la Banque Mondiale (World Bank, 1990) qui stipule que la pauvreté est multidimensionnelle et multi-facette. La définition proposée par la Commission Européenne (*CE*), dans le premier programme d'action pour combattre la pauvreté, synthétise bien cette multidimensionnalité. Pour la *CE*, sont qualifiés de pauvres « des personnes, des familles et des groupes de personnes dont les ressources (matérielles, culturelles, sociales) sont si limitées qu'elles ne leur donnent pas accès à un minimum vital acceptable au sein de la société dans laquelle ils vivent (Dieterlen, 2005) : la pauvreté est donc contextuelle. Une définition de travail adoptée par Payne (2005) dans son ouvrage, décrit tout aussi bien ce caractère multidimensionnel de la pauvreté. Pour paraphraser ce chercheur, un individu ou un ménage est qualifié de pauvre quand il fait face à un manque manifeste de ressources financière, émotionnelle, mentale, spirituelle, physique, de support du système dans lequel il vit, de bonnes relations, de la connaissance et la maîtrise des règles cachées qui modèlent le comportement des groupes sociaux. Cette définition de la pauvreté proposée par Payne (2005) se retrouve sous une formulation différente dans l'ouvrage publié par Shipler (2005) qui a aussi insisté sur les interrelations existant entre ces différentes déficiences. Cette conception de la pauvreté est théorisée par Murray et Herrnstein (1994) qui soutiennent que celle-ci résulte des déficiences des individus eux-mêmes. Cette conceptualisation explicative de la pauvreté est complétée par les travaux de Reich et al. (1973) qui argumentent que la pauvreté est aussi liée au cadre institutionnel environnant établi. Amartya Sen, lui, parle de l'ensemble des « capacités² » dont disposent les gens pour

1. Par biens primaires, Rawls entend les libertés fondamentales, la liberté de mouvement et d'emploi, les pouvoirs et prérogatives liés aux fonctions et postes de responsabilité, le revenu et la richesse et les bases sociales du respect de soi.

2. c'est-à-dire leur capacité à réaliser des choses en utilisant les moyens et les ressources appropriées.

« fonctionner » dans une société donnée.

À l'observation, la pauvreté est associée à la possibilité limitée d'obtenir du crédit, à un faible pouvoir d'achat, à des moyens capabilisants limités, à la vulnérabilité aux risques de diverses catégories, à un accès limité à des biens et services, à la vente de leur force de travail et la location des terres pour la production de nourriture, etc. (EENPI, 2010; Dieterlen, 2005). Dans ce même ordre d'idées, la Banque Mondiale citée par Altmann (2011) caractérise ainsi la pauvreté :

« Poverty is hunger. Poverty is lack of shelter. Poverty is being sick and not being able to see a doctor. Poverty is not being able to go to school and not knowing how to read. Poverty is not having a job, is fear for the future, living one day at a time. Poverty is losing a child to illness brought about by unclean water. Poverty is powerlessness, lack of representation and freedom »

Quand on l'observe et on l'évalue, la pauvreté peut aussi s'apprécier non seulement sur le plan de ses manifestations sociales mais encore spatiales et temporelles. Les disparités régionales, sectorielles et urbain-rural en termes de pauvreté s'avèrent très importantes. Huppi et Ravallion (1991) ont identifié pour l'Indonésie des disparités importantes en termes d'incidence et de sévérité de la pauvreté entre différents secteurs de l'économie sur la période 1984-1987. En particulier, des différences significatives ont été observées entre des secteurs d'emploi en milieu urbain et en milieu rural. De plus, les mesures de pauvreté sont plus élevées en milieu rural qu'en milieu urbain pour un secteur d'emploi donné en particulier entre les fermiers. L'importance et la concentration de la pauvreté en milieu rural comparativement au milieu urbain ont été aussi observées au Liban (Bérenger et al., 2013). Dans le cas haïtien par exemple, cette même tendance à la concentration de la pauvreté en milieu rural a été observée (FMI, 2020). En effet, selon les estimations et prévisions réalisées par Pokhriyal et al. (2020), de très hauts niveaux de pauvreté sont prévus pour des communes situées à l'est du département de l'Artibonite, dans certaines communes des départements du Sud, du Sud-est et du Centre. Ces prévisions restent cohérentes avec les données de la dernière Enquête sur les Conditions des Vie des Ménages Après le Séisme (ECVMAS) réalisée en Haïti par l'Institut Haïtien de Statistique

et d'Informatique (IHSI) en 2012.

Les différentes caractéristiques de la pauvreté précédemment précisées témoignent d'une part de son caractère multidimensionnel, des différents facteurs qui lui sont associés, de la façon dont celle-ci se répartit sectoriellement, géographiquement et en termes urbain/rural. Tout ceci pour signifier l'ampleur du travail à réaliser pour cerner la pauvreté sur le plan conceptuel et la « mesurer », la décrire sur le plan empirique. Dans le cadre de cette thèse sont développés trois essais tentant d'apporter une réponse tout au moins partielle à certains aspects choisis et très précis de la pauvreté étant donné son caractère multidimensionnel et multifacette.

La réussite d'une politique de lutte contre la pauvreté dépend tout d'abord de la capacité à identifier et à dénombrer les pauvres. Les ménages ou individus vivant dans la pauvreté (extrême) sont difficiles à dénombrer du fait qu'ils sont stigmatisés³. Le premier essai aborde en partie cette problématique en mettant au point une procédure de « régression » (économétrique) adaptée à des situations où l'on ne peut pas supposer *a priori* que l'on dispose d'un échantillon représentatif, mais où l'on dispose par contre de *plusieurs* sources de données ou échantillons relatifs à cette même population. Cet essai

3. Selon Baumberg (2016), la stigmatisation personnelle se définit comme le sentiment propre d'une personne pour qui demander est honteux. Ce qui implique que pour cette personne, recevoir une prestation est déshonorant. Du point de vue social, la stigmatisation est un processus englobant des attitudes, des pensées et des actions de la part de la majorité, les perceptions et les réponses à celles-là par les personnes stigmatisées. Cette définition conceptualise la stigmatisation vécue qui constitue un aspect important contribuant à alimenter davantage la stigmatisation en général. Dans leur chapitre d'ouvrage traitant de l'identification et de l'échantillonnage des groupes de population vulnérables et stigmatisés, Berry et Gunn (2014) définissent comme vulnérables et stigmatisés des groupes de population publiquement identifiés et sujets à la honte, au mépris, au ridicule ou à la discrimination dans leurs interactions avec les autres [...]. « Les personnes vivant dans la pauvreté ont honte parce qu'elles transgressent la norme de l'autosuffisance financière. [...] Elles ont honte de leur situation, des autres et de leurs relations avec les institutions prestataires. Des évidences sont établies dans différentes recherches menées sur cette problématique en Grande-Bretagne, en Chine, en Inde, en Norvège, au Pakistan et en Ouganda. » (Walker et Bantebya-Kyomuhendo, 2014).

La difficulté d'identifier et d'échantillonner les groupes de population vulnérables et stigmatisés a été problématisée et documentée par Berry et Gunn (2014). Ces chercheurs ont passé en revue deux approches (identification et échantillonnage de ces groupes comme membres d'un groupe d'intérêt ou en tant que groupes définis par l'objectif spécifique de l'étude) généralement utilisées et des méthodes (échantillonnage boule de neige, échantillonnage sur des sites répertoriés) permettant d'enquêter ces groupes de population. Ils ont répertorié et présenté diverses applications de ces méthodes. Une application intéressante sur cette problématique a été réalisée par Nicaise et Schockaert (2014) à partir des résultats d'une enquête de l'Union Européenne sur le revenu et les conditions de vie. Deux des trois objectifs poursuivis dans cette recherche ont été : a) d'identifier les groupes présentant un risque élevé de pauvreté mais qui n'ont pas été enquêtés ; b) de procéder à une enquête pilote sur les conditions de vie de quelques-uns de ces « groupes oubliés (non observés) de personnes pauvres ».

a aussi comme objectif intermédiaire d'estimer de manière robuste la taille des différents sous-groupes de ces populations, en général sous-enquêtées (ménages pauvres, agriculteurs de subsistance, ...), lorsqu'on procède à leur stratification pour des modélisations et analyses inférentielles.

Les différentes recherches réalisées sur la structure urbain-rural de la pauvreté ont toujours fait ressortir la concentration et la sévérité de la pauvreté rurale. Etant donné l'importance des activités agricoles dans le milieu rural, en particulier dans les pays sous-développés, on développe, dans le cadre du deuxième essai, un modèle pour évaluer les bienfaits et les limites de l'arrosage manuel, sur la base de l'utilité et du revenu espéré des exploitants.

Etant donné leurs conditions de vie déplorables, en particulier la faiblesse de leur revenu, les ménages pauvres ont un faible pouvoir d'achat limitant ainsi leur accès à des biens et services de base. Le troisième essai, dans le cadre d'un modèle simple, étudie les circonstances dans lesquelles une entreprise productrice de biens et services trouverait profitable de ne pas exclure de consommateurs. La problématique générale consistait à essayer de faire le lien entre distribution des revenus au sein de la population, niveau des prix de marché et accès aux biens et services pour les ménages les plus pauvres.

Bibliographie

- ALTMANN, M. P. (2011). *Contextual Development Economics : A Holistic Approach to the Understanding of Economic Activity in Low-Income Countries*. Springer.
- BAUMBERG, B. (2016). The stigma of claiming benefits : a quantitative study. *Journal of Social Policy*, 45(2):181–199.
- BENICOURT, E. (2001). La pauvreté selon le PNUD et la Banque Mondiale. *Études rurales*, (159-160):35–54.
- BÉRENGER, V. et AL. (2013). *Poverty and Social Exclusion around the Mediterranean Sea*, chapitre : Regional and sectoral distributions of poverty in Lebanon, 2004, page 109.
- BERRY, S. H. et GUNN, P. P. (2014). *Hard-to-Survey Populations*, chapitre " Conducting research on vulnerable and stigmatized populations ", page 368. Cambridge University Press.
- DIETERLEN, P. (2005). *Poverty : a philosophical approach*, volume 6. Rodopi.
- EENPI (2010). Markets and Households on Low Incomes Europe Economics and New Policy Institute. Rapport technique, Europe Economics and New Policy Institute (EENPI), Office of Fair Trading.
- FMI (2020). HAITI : SELECTED ISSUES. Rapport technique, Fonds Monétaire International (MFI), Country Report No. 20/122.
- HUPPI, M. et RAVALLION, M. (1991). The sectoral structure of poverty during an adjustment period : evidence for Indonesia in the mid-1980s. *World development*, 19(12):1653–1678.
- MURRAY, C. et HERRNSTEIN, R. J. (1994). Race, Genes and I.Q. : An Apologia. *The New Republic*, 211(18).

- NICAISE, I. et SCHOCKAERT, I. (2014). *Hard-to-Survey-Populations*, chapitre "The hard to reach among the poor in Europe : Lessons from Eurostat's European Union Surveys of Income and Living Conditions in Belgium", page 541. Cambridge University Press.
- PAYNE, R. K. (2005). *A framework for understanding poverty*. Aha! Process Inc.
- POKHRIYAL, N. et AL. (2020). Estimation et prévision de la pauvreté et des inégalités de revenus en Haïti. *Banque Interaméricaine de Développement*.
- REICH, M. et AL. (1973). Dual Labor Markets : A Theory of Labor Market Segmentation. *American Economic Review*, 2(63):359–365.
- SHIPLER, D. K. (2005). *The Working Poor : Invisible in America*. New York, Vintage-Random.
- WALKER, R. et BANTEBYA-KYOMUHENDO, G. (2014). *The Shame of Poverty*. Oxford University Press, USA.
- WORLD BANK (1990). *World development report 1990 : poverty*. Oxford University Press for The World Bank.

Essai # 1

Régression binaire par la méthode de
capture-recapture

Résumé. L'objectif de cet essai est de développer une extension des méthodes dites de « péchage-répêchage » afin de pouvoir procéder à des analyses de populations difficiles à rejoindre. Plus précisément, il s'agit de mettre au point une procédure de « régression » (économétrique) adaptée à des situations où l'on ne peut pas supposer *a priori* que l'on dispose d'un échantillon représentatif, mais où l'on dispose par contre de *plusieurs* sources de données ou échantillons relatifs à cette même population. Ce travail a aussi comme objectif intermédiaire d'estimer de manière robuste la taille des différents sous-groupes de ces populations, en général sous-enquêtées (ménages pauvres, agriculteurs de subsistance, ...), lorsqu'on procède à leur stratification pour des modélisations et analyses inférentielles.

Après avoir passé en revue les principaux estimateurs disponibles à ce jour pour estimer la taille d'une population difficile à rejoindre, nous (re-)calculons l'expression de la variance de l'estimateur de taille de population de Lincoln-Petersen, \hat{n} , pour la comparer à celle de la taille de la population totale observée, o , lorsqu'on procède à plusieurs « sondages ». La (taille de la) population observée o est par nature inférieure ou égale à la (taille de la) population réelle n , dont elle constitue en quelque sorte un estimateur biaisé. Lorsque les sources ou échantillons de données peuvent être supposés indépendants, l'estimateur \hat{n} est un estimateur sans biais de la population totale – et donc *a priori* préférable. Néanmoins, les bénéfices de son adoption dépendent de l'effective indépendance des échantillons ainsi que de leur taille, c'est-à-dire de la probabilité avec laquelle une partie de la population échappe aux observations répétées – qui détermine ces deux variances.

Les effets de sélection associés à la collecte d'information dépendent en général des caractéristiques des individus. La prise en compte de l'hétérogénéité des populations (et donc sa stratification) permet *a priori* d'obtenir des estimateurs « plus fins », c'est-à-dire sans biais. Cependant, elle a aussi un « coût », leur moindre précision, puisque leur variance, en général, est plus élevée. Dans le cas d'un ensemble de caractéristiques finies $z \in Z$, on peut essayer d'estimer la probabilité d'occurrence de chacune d'entre elle au sein de la population. A la différence de l'estimateur de la taille de population o mentionné plus haut, les estimateurs naïfs de ces probabilités $P(z = z_i)$ n'exhibent pas un

biais systématique. Tout dépend de la probabilité des différents individus d'échapper à l'observation. Si cette probabilité pour les individus qui ont certaines caractéristiques est identique à celle observée pour la population dans son ensemble (et donc si l'effet de sélection est identique) ces estimateurs naïfs sont naturellement sans biais.

Finalement, nous considérons le cas où chaque individu de la population est caractérisé par deux variables binaires $x \in \{0, 1\}$ et $y \in \{0, 1\}$. Afin d'estimer les paramètres du modèle binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$, l'expression des probabilités conditionnelles $\bar{q} = P(y_i = 1 | x = 1)$ et $\underline{q} = P(y_i = 1 | x = 0)$ ainsi que leur variance sont calculées. Des estimateurs statistiques de ces grandeurs sont proposés et comparés à leurs correspondants « naïfs », c'est-à-dire calculés à partir des seules observations, sans prendre en compte l'effet de sélection.

Le biais des estimateurs naïfs dépend de la distribution des caractéristiques au sein de la population. Sous certaines conditions, les estimateurs naïfs sont asymptotiquement sans biais - et donc à privilégier par rapport aux estimateurs qui prennent en compte l'hétérogénéité et l'effet de sélection, nécessairement moins précis. C'est pourquoi nous proposons un test statistique de biais nul de l'estimateur naïf.

Lorsque le biais des estimateurs naïfs ne peut être considéré comme étant nul et que la taille des échantillons statistiques est suffisante pour que la prise en compte de l'hétérogénéité puisse être considérée comme bénéfique, les paramètres du modèle

$$P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$$

avec effet de sélection sont estimés. En d'autres termes, nous exhibons la formule des estimateurs

$$\underline{\hat{q}} = P(\widehat{y_i = 1/x = 0}), \widehat{\bar{q}} = P(\widehat{y_i = 1/x = 1}).$$

ainsi que leur variance, tels que calculés en généralisant la logique de Lincoln-Petersen. Par suite, les estimateurs de β_0 et β_1 , pour des raisons statistiques liées fondamentalement au codage arbitraire de y et à la contrainte $0 \leq \text{Prob}(y_i = 1 | x_i) = \beta_0 + \beta_1 x_i \leq 1$, sont

donnés par (Φ étant la fonction de répartition connue de ε_i) :

$$\begin{aligned}\widehat{\beta}_0 &= \Phi^{-1} \left(P(y_i = 1/x = 0) \right) \\ \widehat{\beta}_1 &= \Phi^{-1} \left(P(y_i = 1/x = 1) \right) - \Phi^{-1} \left(P(y_i = 1/x = 0) \right)\end{aligned}$$

1.1 Introduction

Dans de nombreux contextes, l'inexistence de bases de sondage empêche la réalisation d'enquêtes fiables. Cette situation s'observe tant pour des populations animales difficiles à joindre que pour certaines populations humaines difficiles à enquêter. En définissant une population difficile à enquêter comme une population difficile à échantillonner, à identifier, à persuader de prendre part à une enquête, à contacter et difficile à interviewer, Tourangeau et al. (2014) soulèvent les nombreuses difficultés que posent le recensement et l'échantillonnage de ces populations. En épidémiologie en particulier et dans certaines sphères scientifiques (sciences de la vie, sciences médicales), le besoin d'utiliser des méthodes de surveillance efficaces là où il n'est pas possible de disposer d'une base de sondage, se pose en toute acuité. Cette situation est d'autant plus aigüe dans les pays ne disposant pas de ressources importantes pour la constitution et le monitoring de telles bases de sondage. Les méthodes dites de « capture-recapture » dont la problématique est d'estimer la taille d'une population dans un contexte où il n'est pas possible de procéder à un recensement (Rivest, 2013), sont de plus en plus utilisées et permettent de corriger au moins partiellement les biais liés à la non-représentativité des « échantillons » issus des données partielles disponibles relatives à ces populations. Dans leur version multiliste, ces méthodes fournissent des outils permettant d'estimer la taille inconnue des populations humaines difficiles à enquêter à partir de différentes listes de données (listes complètes ou incomplètes) jouant le rôle "d'échantillons". Différents cas d'application tels par exemple l'estimation du nombre de sans-abris, de prostituées, de réfugiés de guerre, d'immigrés clandestins, des individus vivant en extrême pauvreté, etc. sont répertoriés dans la littérature (Chao, 2001). Du nombre de victimes dans une guerre (Manrique-Vallier et al., 2013) ou dans des conflits armés (Mitchell, 2014), au nombre de personnes souffrant d'une pathologie donnée, le nombre d'unités non répertoriées lors d'un recensement (Zwane et Van Der Heijden, 2008), l'usage des méthodes de *CR* ne cesse de prendre de l'ampleur.

Différents outils et méthodes statistiques, entre autres les modèles de régression logistiques et log-linéaires, sont en général utilisés pour développer et améliorer les méthodes d'estimation de la taille de ces populations animales difficiles à joindre et en particulier,

dans la version appliquée en populations humaines, la taille des certaines populations difficiles à enquêter. Il en ressort que l'objectif essentiel de ces méthodes, bien qu'elles utilisent à cette fin des covariables, quand les données disponibles l'autorisent, est *d'estimer la taille* des populations difficiles à joindre ou à enquêter. Nous nous intéressons de manière générale au développement d'une extension de ces méthodes permettant de procéder à une analyse explicative par régression disposant d'échantillons (listes) non représentatifs issus de populations difficiles à joindre et des estimateurs obtenus par les méthodes de pêche répêchage. Pour atteindre cet objectif, on re-calculera l'estimateur de la taille de la population en prenant en compte l'hétérogénéité éventuelle de celle-ci. Ensuite, on définira et caractérisera l'estimateur naïf de la taille de la population observée comparativement à celui de Lincoln-Petersen. Le développement de cette extension se restreindra au cas de deux listes et se portera sur l'estimation d'une relation explicative liant une variable dépendante binaire à une variable indépendante binaire.

Le chapitre sera ainsi organisé. On présentera dans la section 1.2 une vue générale synthétique des méthodes de capture recapture avec une attention particulière pour l'approche multi liste. Dans la section suivante (section 1.3), les principaux résultats du modèle seront précisés. On décrira le modèle à la section 1.4. Cette section commence par la définition des principales notations qui sont utilisées. Après avoir décrit l'historique de capture pour deux listes, seront précisés les estimateurs associés pour estimer la taille d'une population. On définira et caractérisera l'estimateur naïf de la taille de la population et on évaluera l'opportunité de l'utiliser ou non pour estimer les paramètres de notre modèle de régression exprimant une variable dépendante binaire en fonction d'une variable indépendante binaire. A la section 1.5, on analysera les résultats par rapport à la littérature existant. Enfin, on présentera quelques remarques conclusives.

1.2 Méthodes capture-recapture (CR)

1.2.1 Définition, but, origines, méthodes et modèles

1.2.1.1 Définition et but

Les méthodes de capture-recapture (pêchage-répêchage, contact-recontact) constituent un ensemble de méthodes servant à décrire et à estimer les paramètres liés à l'état d'une population et les processus qui en sous-tendent la dynamique en absence de base de sondage et/ou lorsque la population est observée de manière imparfaite. Cette définition, adaptée de celle proposée par Vergne (2012), donne déjà une indication sur le but fondamental poursuivi à travers le développement des méthodes de capture-recapture (CR). En effet, ces méthodes se donnent fondamentalement pour objectifs de : i) caractériser une population en termes d'abondance et de composition ; ii) étudier les variations temporelles d'une population en termes d'abondance et de composition et aussi iii) décrire et estimer les processus liés à la dynamique d'une population (*ibid*).

Pour atteindre les objectifs poursuivis, les différents modèles développés dans le cadre de ces méthodes prennent en compte l'ensemble des caractéristiques distinctives et dynamiques des populations sous-étude. Dans une démarche de *CR*, les populations d'intérêt sont étudiées et modélisées suivant deux hypothèses fondamentales. La première hypothèse adoptée est celle de *population fermée*. Cette hypothèse est utilisée dans le cas où il est plausible de supposer qu'il y a pas d'entrée (naissances, immigration) ni de sortie (décès, émigration) entre deux occasions de capture au sein de la population sous-étude. La deuxième hypothèse est celle de *population ouverte*. Cette hypothèse est envisagée dans le cas où la dynamique des populations étudiées produit des changements non négligeables qui doivent être pris en compte. D'autres hypothèses spécifiques sont envisagées dans le cas des modèles spécifiques.

1.2.1.2 Les origines de ces méthodes

Les méthodes dites de « capture-recapture » (pêchage-répêchage, contact-recontact), formellement initiées par les travaux de C. G. J. Petersen en 1836 et de F. C. Lincoln

en 1930, ont été développées en biologie afin d'estimer la taille de populations animales difficiles à joindre. Petersen s'était intéressé à l'estimation de la taille d'une population de poissons et Lincoln à l'estimation de la taille d'une population de sauvagines (oiseaux aquatiques sauvages tels canards, oies ou échassiers). Les origines des premières réflexions et utilisations de ces méthodes remontent un peu plus loin dans le temps si on s'en tient aux travaux de Laplace en 1786 qui utilisa ces méthodes pour estimer l'effectif de la population française et aux travaux de John Graunt en 1600 pour apprécier la structure de la population du Royaume Uni (Le Cren, 1965; Lavallée et Rivest, 2013; Vergne, 2012).

La démarche de ces méthodes consiste, au moyen de "pièges" appropriés, à capturer un sous-groupe d'animaux qui composent la population, de les marquer puis de les relâcher après chaque "occasion" de capture. Cet exercice est répété de manière indépendante plusieurs fois (intervalles réguliers) ou "occasions". Si l'opération est menée à K reprises (occasions) ($K > 1$), un membre de cette population peut être capturé K fois ou $K-1$ fois, ..., ou encore jamais. Plus précisément, il peut être capturé lors de la première "occasion" de capture – ou non, lors de la deuxième – ou non, et ainsi de suite jusqu'à la dernière occasion. A chaque membre de la population, on peut donc associer un « historique de capture ». Il existe au total 2^K historiques de capture possibles. Les méthodes de capture-recapture utilisent les informations fournies par les $2^K - 1$ historiques de capture qui donnent lieu à observation pour estimer le nombre d'animaux qui ne sont jamais observés et, in fine, la taille de la population totale.

1.2.1.3 Méthodes et modèles

Les méthodes utilisées dans la démarche pêcheage répêcheage dépendent des caractéristiques de mouvement au sein de la population sous-étude. En général, les chercheurs formulent des hypothèses plausibles quant au mouvement réel observable au niveau de la population étudiée. De ce fait, ils élaborent deux types de méthodes. Une première méthode se développe sur la base de l'hypothèse de population fermée et l'autre méthode suppose l'existence de mouvements (population ouverte) au sein de la population étudiée. Sont construits pour chaque méthode des modèles ou classe de modèles spécifiques

reposant sous des hypothèses appropriées.

L'hypothèse de population fermée postule l'absence de mouvements au sein de la population étudiée. Autrement dit, cette hypothèse sous-tend que sur une période donnée (relativement courte), l'ensemble des entrées (naissance, immigration, ...) et l'ensemble des sorties (décès, émigration, ...) au sein de la population sont sensiblement égaux de sorte que la taille N de cette population est supposée inchangée. Différentes méthodes d'estimation sont utilisées pour estimer la taille N d'une population sous cette hypothèse. On y trouve les méthodes d'estimation de la densité d'une population et les méthodes par capture recapture. Dans le cas de l'estimation de la densité d'une populations, on dispose des méthodes relevant des unités d'échantillonnage telles entre autres, le recensement aérien, les méthodes de distance, la méthode d'enlèvement et la méthode de densité relative (Seber, 1986). Les méthodes par capture recapture opèrent sous l'hypothèse de captures uniques ou de captures multiples.

Dans le cas des méthodes par captures uniques, en supposant qu'une population de taille inconnue N répartie en \mathcal{U} éléments jamais capturés et \mathcal{M} éléments qui ont été déjà capturés, *également répartis dans la population*, on choisit *un échantillon aléatoire* de taille n réparti en u éléments jamais capturés et m éléments qui ont été capturés. Les estimateurs proposés par Petersen et leurs extensions sont mis à profit dans ce cas de figure. Les méthodes par captures multiples ou recensement de Schnabel (Schnabel, 1938) consiste à choisir une série d'échantillons aléatoires avec remise en procédant de façon similaire aux méthodes par captures uniques. Différents modèles, en raison d'hypothèses diverses, sont proposés pour estimer la taille inconnue N d'une population par captures multiples. On y trouve les modèles loglinéaires (Heckel et Roughgarden, 1979) et une classe de modèles prenant en compte les variations dans la capturabilité des animaux (Otis *et al.*, 1978). Le modèle développé par Heckel et Roughgarden (1979) intègre la technique d'échantillonnage par cadrat¹ couplée à une procédure de collecte de données sur le terrain à partir d'un recensement.

Il existe d'autres classes de modèles tenant compte des variations dans la capturabilité

1. Cette technique d'échantillonnage consiste à définir et à circonscrire une zone (une aire géographique, un quadrat) déterminée et à recenser l'ensemble des animaux qui s'y trouvent.

des animaux . On y recense sept types de modèles M_j où $j \in \{o, b, h, t, bh, th, tbh\}$. En admettant que chaque animal a la même probabilité d'être attrapée, ces modèles estiment la taille inconnue N de la population en supposant inchangé le comportement des animaux (o) ou en tenant compte de leur comportement face au piège (b) et/ou de leur degré d'hétérogénéité (h) et/ou de la variabilité des probabilités de capture intra-échantillon (t).

Dans le cas des populations supposées ouvertes, le modèle de Jolly-Seber (JS) s'inscrit dans le cadre des méthodes par captures multiples et constitue l'un des modèles fondateurs prenant en compte les décès, la migration permanente et les immigrations (Seber et Manly, 1985). D'autres chercheurs ont élaboré des extensions du modèle JS (Buckland *et al.*, 1983). D'autres modèles développés sous le nom de modèles de récupération d'étiquettes d'animaux décédés considèrent une cohorte d'individus marqués et libérés chaque année. Les étiquettes sont récupérées à partir d'animaux trouvés morts ou exploités (Seber, 1985).

La version multiliste, adaptée des méthodes CR sous l'hypothèse de population fermée, est très utilisée de nos jours pour estimer la taille des populations humaines difficiles à enquêter. L'usage de ces méthodes tient son origine de leurs applications en épidémiologie et d'autres sciences connexes. L'un des pionniers dans l'usage de ces méthodes fut Sekar et Deming (1949) qui, en exploitant deux listes disponibles (dual system, dual record system), avaient appliqué ces méthodes à l'estimation des taux de naissance et de décès en Inde entre 1945 et 1946. Les chercheurs Wittes et Sidel (1968) ont été les premiers à utiliser « a three sample records » pour estimer le nombre de patients d'un hôpital (Chao, 2014) et à généraliser les résultats pour un nombre de k listes (multisystem records).

L'extension que nous développons s'inscrit dans la version multiliste des méthodes de CR . Dans la prochaine section, on présentera très succinctement le cadre général, les principales hypothèses et les principaux éléments caractéristiques de la version multiliste de ces méthodes.

1.2.2 Approche multiliste

Dans le cas de populations humaines, les méthodes dites de « capture-recapture » s'appliquent sur des listes partielles considérées comme des échantillons (non représentatifs) de la population à étudier. Le problème fondamental de cette approche consiste donc à estimer la taille d'une population à partir d'un ensemble de listes incomplètes partielles disponibles - l'approche multi-liste - dans un contexte où il n'est pas possible de procéder à un recensement (Lavallée et Rivest, 2013; Otis *et al.*, 1978) . Ces méthodes sont utilisées dans divers champs disciplinaires des sciences sociales et des sciences médicales et de la vie (IWGD MF, 1995) pour inférer, à partir des listes (incomplètes) existantes sur la taille et quand les données listées le permettent, sur les principales caractéristiques de certaines populations difficiles à enquêter. Les méthodes *CR* fournissent des outils et techniques permettant d'estimer par exemple, le nombre de sans abris, l'importance du nombre des enfants travailleurs, la prostitution infantile, l'estimation du nombre de naissances et de décès pendant l'année, l'estimation du nombre d'unités non répertoriées lors d'un recensement, etc. En sciences médicales et sciences de la vie, ces méthodes sont utilisées pour estimer la prévalence de certaines maladies cachées telles par exemple la dépendance à la drogue, la prévalence du VIH, pour évaluer les systèmes épidémiologiques de surveillance, etc. (Chao, 2014; Zwane et Van Der Heijden, 2008; Chao, 2001) .

Les principales caractéristiques de l'approche multi-liste sont décrites et précisées dans Chao (2014) :

« Le cadre de modélisation des systèmes à enregistrements multiples pour les populations humaines est similaire à celui d'un système fermé tel que supposé dans les modèles de capture-recapture pour l'estimation des populations animales : chaque liste est considérée comme un échantillon non représentatif issu de captures et les numéros d'identification et / ou les noms sont utilisés comme une étiquette. La « capture dans un échantillon » correspond à « ayant été enregistrée ou identifiée dans une liste » et « probabilité de capture » devient « probabilité d'enregistrement ». Sont à signaler deux grandes différences entre les applications dans le cas des populations animales et celles des populations humaines :

- i) Il existe plus d'échantillons issus des occasions de capture dans le cas des populations

animales que de listes disponibles dans le cas des populations humaines et (ii) dans les études sur les populations animales, il existe un ordre naturel chronologique ou séquentiel dans les échantillons par capture recapture, alors que pour les données épidémiologiques, cet ordre n'existe pas dans les listes, ou peut être différent pour certaines personnes ».

On peut synthétiser l'historique d'observations (capture) des N individus sur K listes (occasions de capture) de la manière suivante. Soit alors la variable Y_{ik} définie par :

$$Y_{ik} = \begin{cases} 1, & \text{si l'individu } n^oi \text{ a été observé (capturé) dans la liste } n^ok \\ 0, & \text{si l'individu } n^oi \text{ n'a pas été observé dans la liste } n^ok \end{cases}$$

L'historique d'observation H des N individus sur les K listes (K occasions de capture) est noté :

TABLE 1.1 – Historique de capture de N individus sur K listes
$$H = \begin{array}{c|cccccc} \nearrow & 1 \dots & 2 & \dots & k & \dots & K \\ \hline 1 & Y_{11} & Y_{12} & \dots & Y_{1k} & \dots & Y_{1K} \\ 2 & Y_{21} & Y_{22} & \dots & Y_{2k} & \dots & Y_{2K} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ i & Y_{i1} & Y_{i2} & \dots & Y_{ik} & \dots & Y_{iK} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ N & Y_{N1} & Y_{N2} & \dots & Y_{Nk} & \dots & Y_{NK} \end{array}$$

Désignons par $Z_{s_1 s_2 \dots s_K}$ le nombre d'individus ayant un historique de capture s_1, s_2, \dots, s_K où

$$s_k = \begin{cases} 1, & \text{pour tout individu } i \text{ présent dans la liste } n^ok \\ 0, & \text{pour tout individu } i \text{ absent de la liste } n^ok \end{cases}.$$

Le nombre d'individus ayant un historique de capture s_1, s_2, \dots, s_K est défini par (Tsay et Chao, 2001; Chao et Tsay, 1998) :

$$Z_{s_1 s_2 \dots s_K} = \sum_{i=1}^N I_{[Y_{i1}=s_1, \dots, Y_{iK}=s_K]}$$

où Y_{ik} , $k = [1, K]$, correspond à la ligne i du tableau 1.1 et $I_{[A]}$ au nombre de fois qu'un événement A se réalise. Pour un indice k fixé tel que l'indice $n^o s_k = 1$, il y a un nombre 2^{K-1} de $Z_{s_1 s_2 \dots s_K}$ obtenus par *permutation avec remise des indices* ($s_1, \dots, s_k = 1, \dots, s_K$). Le nombre total N_k d'individus observés dans la liste $n^o k$, $k = [1, K]$ correspond alors à la somme des $Z_{s_1 s_2 \dots s_K}$ obtenus par permutation avec remise.

Exemple. Pour $K = 4$ occasions de capture, on a une série de $Z_{s_1, \dots, s_k=1, \dots, s_4}$ correspondant au nombre de $2^{4-1} = 8$.

— Pour $k = 1, s_1 = 1$. On obtient : $Z_{1000}, Z_{1100}, Z_{1110}, Z_{1111}, Z_{1001}, Z_{1010}, Z_{1101}, Z_{1011}$.

Dans ce cas,

$$N_1 = Z_{1000} + Z_{1100} + Z_{1110} + Z_{1111} + Z_{1001} + Z_{1010} + Z_{1101} + Z_{1011};$$

— Pour $k = 2, s_2 = 1$. On a : $Z_{1100}, Z_{0100}, Z_{1110}, Z_{1111}, Z_{1101}, Z_{0111}, Z_{0101}, Z_{0110}$. Par suite,

$$N_2 = Z_{1100} + Z_{0100} + Z_{1110} + Z_{1111} + Z_{1101} + Z_{0111} + Z_{0101} + Z_{0110};$$

et ainsi de suite.

Toujours dans l'objectif d'améliorer les méthodes disponibles pour estimer la taille d'une population humaine par l'approche multiliste, Tsay et Chao (2001); Chao et Tsay (1998) ont développé une méthode dénommée « couverture échantillonnale ² » (sample coverage) qui s'interprète comme une mesure de chevauchement informationnel entre plusieurs enregistrements dans les différentes listes. Cette méthode permet d'estimer la taille d'une population humaine dont un échantillon est enregistré dans K listes en prenant en compte les dépendances existant entre les paires de listes sous un ensemble d'hypothèses appropriées.

Les progrès réalisés dans l'élaboration et l'application des méthodes statistiques dites de « capture-recapture » vont presque toujours dans le sens d'une amélioration des mé-

2. La couverture échantillonnale est définie comme la fraction, à probabilité pondérée, de la population récupérée dans l'échantillon.

thodes d'estimation de la taille d'une population là où il y a absence de base de sondage. Pour un nombre d'individus n^{yx} de caractéristiques (x, y) , tout en tenant compte de l'hétérogénéité des membres de la population, on se propose, *dans une démarche de régression*, estimer directement, à partir des estimateurs fournis par les méthodes *CR* le modèle de régression binaire $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$. En d'autres termes, en estimant la probabilité d'observer le nombre d'individus de caractéristique y étant donné une caractéristique x . Le modèle est développé dans le cas de la disponibilité deux listes. Une caractérisation du biais de l'estimateur naïf de cette probabilité par rapport à l'estimateur obtenu par les méthodes de *CR* est présentée. Étant donné la dépendance non linéaire de cet estimateur de la fréquence non observée pour des individus de caractéristique quelconque yx , une approximation de l'estimateur de la variance de l'estimateur est obtenu par la Méthode Delta en supposant qu'il y a indépendance entre les observations faites sur les deux listes.

1.3 Les principaux résultats du modèle

L'objectif de ce chapitre consiste fondamentalement à initier une analyse de régression dans le cadre des estimateurs fournis par les méthodes de capture recapture. Les paramètres du modèle binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$, $x_i, y_i \in \{0, 1\}$, sont estimés à partir des estimateurs \hat{q} et \hat{q} de Lincoln-Petersen développés pour estimer la taille d'une population difficile à joindre. Les estimateurs $\hat{\beta}_0$ et $\hat{\beta}_1$ obtenus dans le cadre des estimateurs fournis par les méthodes de capture-recapture - *EMCR* - pour β_0 et β_1 s'écrivent (Φ étant la fonction de répartition connue de ε_i) :

$$\begin{aligned}\hat{\beta}_0 &= \Phi^{-1}(\hat{q}) \equiv \Phi^{-1}\left(\frac{\widehat{n^{10}}}{\widehat{n^{10}} + \widehat{n^{00}}}\right) \\ \hat{\beta}_1 &= \Phi^{-1}(\hat{q}) - \Phi^{-1}(\hat{q}) \\ &= \Phi^{-1}\left(\frac{\widehat{n^{11}}}{\widehat{n^{11}} + \widehat{n^{01}}}\right) - \Phi^{-1}\left(\frac{\widehat{n^{10}}}{\widehat{n^{10}} + \widehat{n^{00}}}\right)\end{aligned}$$

où

$$\begin{aligned}\widehat{q} &= P(y_i = 1/x_i = 0) \\ \widehat{q} &= P(y_i = 1/x_i = 1).\end{aligned}$$

En maintenant l'hypothèse d'indépendance des observations au niveau des deux sources et en relaxant l'hypothèse de constance de $n_{11}^{yx} + n_{10}^{yx}$, $n_{11}^{yx} + n_{01}^{yx}$, les estimateurs \widehat{q} et \widehat{q} ont été comparés à leur correspondant naïf respectif \widetilde{q} et \widetilde{q} . Pour calculer la variance des estimateurs \widehat{q} et \widehat{q} (basés sur ceux de Lincoln-Petersen) et \widetilde{q} et \widetilde{q} (estimateurs naïfs), la variance exacte de $o^{yx} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx}$ a été calculée à partir d'une distribution trinominale conditionnelle et la variance de $\widehat{n}^{yx} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} + \widehat{n}_{00}^{yx}$ a été re-calculée. Nous avons pu démontrer que même asymptotiquement, les estimateurs naïfs sont biaisés et que leur variance est inférieure à celle des *EMCR*.

1.4 Le modèle de Régression Binaire par la Méthode « Capture-Récapture »

Comme on l'avait précisé au niveau de l'introduction, l'objectif essentiel de ce chapitre est de développer une extension des méthodes *CR* permettant de procéder à une analyse explicative par régression disposant d'échantillons (listes) non représentatifs issus de populations difficiles à joindre et des estimateurs obtenus par les méthodes de pêche répêchage. Le développement de cette extension se restreindra au cas de deux listes et à l'estimation d'une relation explicative liant une variable dépendante binaire à une variable indépendante binaire.

En disposant d'échantillons (listes) non représentatifs issus de populations difficiles à joindre et des estimateurs obtenus par les méthodes de pêche répêchage, nous nous intéressons à l'estimation du modèle de régression binaire suivante :

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{1.1}$$

où $\varepsilon_i \sim iid(0, \sigma_\varepsilon^2)$, x_i et y_i sont des variables binaires c'est-à-dire $x_i, y_i \in \{0, 1\}$. Ce qui revient à dire :

$$\begin{aligned} E(y_i / x_i) &= \beta_0 + \beta_1 x_i \\ &= Prob(y_i = 1 / x_i) \text{ car } y_i \sim B(Prob(y_i = 1 / x_i)) \end{aligned}$$

de sorte que le modèle (1.1) se re-écrit de la manière suivante :

$$y_i = Prob(y_i = 1 / x_i) + \varepsilon_i \quad \text{ou mieux} \quad Prob(y_i = 1) = Prob(y_i = 1 / x_i) + \varepsilon_i$$

car $0 \leq Prob(y_i = 1 / x_i) = \beta_0 + \beta_1 x_i \leq 1$.

Pour des raisons statistiques liées fondamentalement au codage arbitraire de y et à la contrainte

$$0 \leq Prob(y_i = 1 / x_i) = \beta_0 + \beta_1 x_i \leq 1,$$

le modèle (1.1) se re-écrit comme suit :

$$y_i^* = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{1.2}$$

où y_i^* est une variable inobservable qualifiée de latente³ explicitant les conditions dans lesquelles $y_i = 1$. Ainsi,

$$y_i = \begin{cases} 1 & \text{si } y_i^* > 0 \\ 0 & \text{sinon} \end{cases} . \tag{1.3}$$

Par suite, $Prob(y_i = 1 / x_i) = Prob(y_i^* > 0) \implies Prob(y_i = 1 / x_i) = Prob(\varepsilon_i > -(\beta_0 + \beta_1 x_i))$.

Donc, $p_i \equiv Prob(y_i = 1 / x_i) = \Phi(\beta_0 + \beta_1 x_i)$ où $\Phi(\cdot)$ ⁴ est la fonction de répartition de la variable aléatoire ε_i de moyenne nulle.

L'objectif est d'estimer, disposant de deux listes jouant le rôle d'échantillons (non

3. Variable (aléatoire) quantitative auxiliaire ou seuil de tolérance dépendant des conditions d'observation de x_i (Gourieroux, 1989).

4. En faisant l'hypothèse de symétrie de la loi de ε_i , on a : $\Phi(x) + \Phi(-x) = 1$. Habituellement, on suppose que les ε_i sont distribuées selon une loi normale centrée réduite ou selon une loi logistique. Cette dernière est surtout utilisée pour des raisons de flexibilité mathématique et de facilité de calcul d'effets marginaux. On rappelle que si $\varepsilon_i \rightarrow iid\mathcal{N}(0, \sigma_\varepsilon^2) \implies \left(\frac{\pi\varepsilon_i}{\sigma_\varepsilon\sqrt{3}}\right) \rightarrow l\left(0, \frac{\pi^2}{3}\right)$.

représentatifs), les paramètres β_0 et β_1 du modèle $p_i = \Phi(\beta_0 + \beta_1 x_i)$ ou de manière équivalente

$$g(p_i) \equiv \Phi^{-1}[\text{Prob}(y_i = 1/x_i)] = \beta_0 + \beta_1 x_i \quad (1.4)$$

tels que⁵

$$\begin{cases} \Phi^{-1}[\text{Prob}(y_i = 1/x_i = 0)] & = \beta_0 \\ \Phi^{-1}[\text{Prob}(y_i = 1/x_i = 1)] - \Phi^{-1}[\text{Prob}(y_i = 1/x_i = 0)] & = \beta_1 \end{cases} \quad (1.5)$$

La description du modèle se structure de la manière suivante. On commencera par préciser en 1.4.1 les principales notations qui sont utilisées. Ensuite, on développera le modèle dans le cas de deux listes dans la sous-section 1.4.2. On y déterminera les estimateurs d'intérêt et leur variance; la variance obtenue à partir de la distribution sous-jacente au processus générateur des données sera comparée à la variance asymptotique utilisée dans la littérature. Ensuite, on définira et évaluera l'estimateur naïf relatif aux estimateurs obtenus par les méthodes de *CR*. Une caractérisation et un test de biais nul seront proposés. Enfin, on présentera quelques simulations numériques. En troisième lieu, on discutera les résultats du modèle eu égard à la littérature. En dernier lieu viennent les remarques conclusives et les perspectives.

1.4.1 Notations générales et définition

En accord avec l'objectif poursuivi, on se propose, pour une variable dépendante y et une variable indépendante x binaires, d'estimer le modèle de régression binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$ re-écrite à l'équation (1.2) à partir d'une règle statistique de décision définie en (1.3). On notera $\bar{q} = P(y_i = 1/x = 1)$ et $\underline{q} = P(y_i = 1/x = 0)$. L'inférence est basée fondamentalement sur les probabilités conditionnelles \bar{q} et \underline{q} et leur caractérisation.

On considère à cet effet une population \mathcal{U} définie par $\mathcal{U} = \{1, 2, \dots, N\}$, N étant la

5. Soit Φ_X la fonction de répartition d'une variable aléatoire réelle X définie par $\Phi_X : \mathbb{R} \rightarrow [0; 1]$. La fonction Φ_X possède les propriétés suivantes : a) Φ_X est non décroissante; b) Φ_X est continue à droite; c) $\lim_{x \rightarrow -\infty} \Phi_X(x) = 0$, $\lim_{x \rightarrow +\infty} \Phi_X(x) = 1$. Soit G une fonction définie sur l'intervalle $]0, 1[$ telle que $G(w) = \inf \{x \in \mathbb{R} / \Phi_X(x) \geq w\}$. Alors la fonction G est une variable aléatoire réelle définie sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ que Φ . La fonction G ainsi définie est la fonction de répartition réciproque de Φ ($\Phi^{-1} \equiv G$) sur $(\Omega, \mathcal{A}, \mathbb{P})$.

taille de la population. On s'intéresse au cas où chaque individu de la population est caractérisé par deux variables binaires $x \in \{0, 1\}$ et $y \in \{0, 1\}$. On note n^{yx} le nombre total d'individus de caractéristique (x, y) .

On suppose qu'on fait K observations partielles (listes, sources) en général de cette population. Une observation (liste) est indiquée par $k = 1, \dots, K$. Pour chaque individu $i \in \mathcal{U}$, on note $\omega^i = (\omega_1^i, \omega_2^i, \dots, \omega_K^i)$ l'historique des captures de l'individu i tel que :

$$\omega_k^i = \begin{cases} 1, & \text{si l'individu apparait dans la liste } n^{\circ k} \\ 0, & \text{sinon} \end{cases}.$$

On désigne respectivement par n_{ω}^{yx} et μ_{ω}^{yx} les fréquences observées et théoriques des individus de caractéristiques (y, x) avec l'historique de capture $\omega \in \Omega = \{00, 01, 10, 11\}$. Le nombre total n^{xy} d'individus de caractéristique (x, y) est donné par :

$$n^{yx} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} + n_{00}^{yx} = \sum_{\omega} n_{\omega}^{yx}.$$

Pour $yx \in \{00, 01, 10, 11\}$, on définit :

$$\begin{aligned} n^{11} &= \{\#i \in \mathcal{U} \mid (y = 1) \wedge (x = 1)\}, \\ n^{10} &= \{\#i \in \mathcal{U} \mid (y = 1) \wedge (x = 0)\}, \\ n^{01} &= \{\#i \in \mathcal{U} \mid (y = 0) \wedge (x = 1)\}, \\ n^{00} &= \{\#i \in \mathcal{U} \mid (y = 0) \wedge (x = 0)\}. \end{aligned}$$

Par suite,

$$N = \sum_{yx} n^{yx}.$$

On se propose d'estimer les probabilités

$$\bar{q} = P\{y = 1 \mid x = 1\} = \frac{n^{11}}{n^{11} + n^{01}}, \quad (1.6)$$

$$\underline{q} = P\{y = 1 \mid x = 0\} = \frac{n^{10}}{n^{10} + n^{00}}. \quad (1.7)$$

On définit

$$\hat{q} = \frac{n^{11} + n^{01}}{N}, \quad (1.8)$$

\hat{q} étant la proportion observée d'individus de caractéristique x de la population. Les trois probabilités \bar{q} , \underline{q} et \hat{q} , définissent la distribution statistique de la population. Le processus générateur des données n^{yx} est une distribution multinomiale :

$$(n^{00}, n^{01}, n^{10}, n^{11}) \sim \mu(N, q^{00}, q^{01}, q^{10}, q^{11}) \quad (1.9)$$

où

$$\begin{aligned} q^{11} &= \frac{n^{11}}{N} = \bar{q}\hat{q}, & q^{00} &= \frac{n^{00}}{N} = (1 - \hat{q})(1 - \underline{q}), \\ q^{10} &= \frac{n^{10}}{N} = \underline{q}(1 - \hat{q}), & q^{01} &= \frac{n^{01}}{N} = (1 - \bar{q})\hat{q}. \end{aligned} \quad (1.10)$$

En supposant que n^{yx} est strictement positif, on définit les probabilités p_ω^{yx} par

$$\mu_\omega^{yx} = n^{yx} p_\omega^{yx},$$

où μ_ω^{yx} est la fréquence théorique et $n^{yx} = Nq^{yx}$ la fréquence observée. Par suite,

$$\mu_\omega^{yx} = Nq^{yx} p_\omega^{yx}.$$

La fonction de vraisemblance s'écrit donc

$$L = \frac{N!}{\prod_{(yx \times \omega)} n_\omega^{yx}!} \prod_{(yx \times \omega)} (q^{yx} p_\omega^{yx})^{n_\omega^{yx}}$$

et la *log-vraisemblance*, à une constante près, a pour expression :

$$\mathcal{L} = \sum_{yx \times \omega} n_\omega^{yx} \ln(q^{yx} p_\omega^{yx}) \quad (1.11)$$

avec $p_{\omega=00}^{yx} = 1 - \sum_{\omega \neq 00} p_\omega^{yx}$.

1.4.2 Cas de deux listes

L'usage de deux listes ou système d'enregistrement dual pour des estimations en populations humaines basées sur les méthodes de capture recapture est très répandu. Un nombre significatif de recherches empiriques en épidémiologie en général et en particulier dans l'évaluation de la couverture d'un recensement disposant d'une enquête postcensitaire, sont réalisées sur la base de l'exploitation de deux sources (listes) (Chatterjee, 2016; Brittain et Bohning, 2009; Zaslavsky et Wolfgang, 1993). L'usage de plus de deux sources est relativement difficile à mettre en oeuvre du fait qu'il est un peu coûteux et que la probabilité d'observation des individus est relativement plus élevée que la probabilité d'observation des animaux (Chatterjee, 2016).

Pour l'usage de deux sources à des fins d'estimation de la taille d'une population difficile à enquêter, les hypothèses de base formulées se résument fondamentalement à trois. Les deux sources sont d'abord supposées indépendantes ; ensuite, il est postulé que la population sous étude est fermée sur la période où la recherche est réalisée ; enfin, on suppose que les individus enregistrés dans les deux listes sont identifiables, appariables et indépendants. Dans certains cas, ces hypothèses sont difficiles d'application et conduisent à un biais positif ou négatif de corrélation ⁶. Cependant, dans le cas de deux sources, un problème de sous-identification ⁷ (Chao et Tsay, 1998) conduit à adopter l'hypothèse d'indépendance bien que cette hypothèse peut se révéler parfois non plausible.

1.4.2.1 Expression matricielle des données de capture recapture pour deux sources

On note par r_1^{yx} la probabilité qu'un individu de caractéristique yx soit présent dans la liste 1 et r_2^{yx} sa probabilité de présence dans la liste 2. Pour chaque individu de caractéristique yx , il est supposé que la probabilité qu'il soit présent ou non dans la première

6. Le problème de biais de corrélation a été défini et caractérisé par Chao et Chiang (2008); Chao et Tsay (1998); Chao et al. (2001). Dans le cas de deux listes, le biais de corrélation peut provenir de deux sources : la dépendance entre la probabilité de présence d'un individu dans une liste et sa présence dans d'autres listes et l'hétérogénéité des individus. L'hypothèse d'indépendance entre les deux sources, quand elle est vérifiée, implique l'indépendance entre les individus. Ces chercheurs ont proposé une méthode d'évaluation et de correction de ce biais de corrélation.

7. L'estimation du biais de corrélation γ_{12} dans le cas de deux sources implique l'estimation simultanée de quatre paramètres μ_1 , μ_2 , N et γ_{12} pour trois cellules de données disponibles.

liste est *indépendante* de sa probabilité d'être présent ou non dans la deuxième. De plus, les deux sources sont aussi supposées indépendantes et qu'il n'y a pas eu ni d'ajout ni de retrait des listes au moment de la recherche. La situation est décrite dans le tableau suivant :

TABLE 1.2 – Matrice des résultats de présence/absence pour deux listes

	Présence dans la liste 2	Absence de la liste 2
Présence dans la liste 1	$n_{11}^{yx} = n^{yx} (r_1^{yx}) (r_2^{yx})$	$n_{10}^{yx} = n^{yx} (r_1^{yx}) (1 - r_2^{yx})$
Absence de la liste 1	$n_{01}^{yx} = n^{yx} (1 - r_1^{yx}) (r_2^{yx})$	$n_{00}^{yx} = n^{yx} (1 - r_1^{yx}) (1 - r_2^{yx})$

où :

- n_{11}^{yx} indique l'ensemble des individus de caractéristique yx observés dans les listes 1 et 2,
- n_{10}^{yx} indique l'ensemble des individus de caractéristique yx observés seulement dans la liste 1,
- n_{01}^{yx} indique l'ensemble des individus de caractéristique yx observés seulement dans la liste 2,
- n_{00}^{yx} indique l'ensemble des individus de caractéristique yx non observés.

L'estimation du total des individus non observés de caractéristique yx et de l'effectif total n^{yx} et l'estimation des probabilités r_1^{yx} et r_2^{yx} permettront d'estimer les probabilités conditionnelles définies en (1.6) et (1.7).

1.4.2.2 Les principaux estimateurs de la taille N dans le cas de deux listes

Pour estimer la taille N d'une population dans le cadre d'un système d'enregistrement dual, différents chercheurs ont proposé des estimateurs élaborés à partir de la disponibilité de deux échantillons (non représentatifs) issus de deux sources indépendantes telles que décrites dans le tableau 1.2. Dans le cadre d'une estimation faite à partir de deux sources, on compte principalement les estimateurs proposés par Lincoln-Petersen, Chapman, Zellerman et Chao.

Estimateur de Lincoln-Petersen. En référence au tableau 1.2 et sous l'hypothèse d'indépendance, se vérifie l'égalité $n_{11}^{yx} n_{00}^{yx} = n_{01}^{yx} n_{10}^{yx}$. Comme les effectifs n_{11}^{yx} , n_{10}^{yx} , n_{01}^{yx} sont observés dans au moins une des deux listes, on en déduit un estimateur de n_{00}^{yx} donné par :

$$\widehat{n_{00}^{yx}} = \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}. \quad (1.12)$$

Par suite, un estimateur de la taille de la population totale n^{yx} de caractéristique yx s'écrit (Annexe 1.A.1)

$$\begin{aligned} \widehat{n^{yx}} &= n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx} + \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} \\ &= \frac{1}{n_{11}^{yx}} (n_{01}^{yx} + n_{11}^{yx}) (n_{10}^{yx} + n_{11}^{yx}). \end{aligned} \quad (1.13)$$

Une variance asymptotique de cet estimateur, déterminée par Sekar et Deming (1949), a pour expression (Annexe 1.A.2) :

$$\begin{aligned} V(\widehat{n^{yx}}) &= \widehat{n_{00}^{yx}} + \left(\widehat{n_{00}^{yx}}\right)^2 \left(\frac{1}{n_{10}^{yx}} + \frac{1}{n_{01}^{yx}} + \frac{1}{n_{11}^{yx}}\right). \\ &= \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{11}^{yx}}\right) \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{11}^{yx}}\right) \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \end{aligned} \quad (1.14)$$

Ces chercheurs ont été les premiers à utiliser l'estimateur de Lincoln-Petersen. L'existence potentielle d'un biais de corrélation entre les deux sources les ont conduits en 1981 à proposer un estimateur d'un coefficient de corrélation défini par :

$$\widehat{\rho^{CD}} = \frac{n_{11}^{yx} \widehat{n_{00}^{yx}} - n_{01}^{yx} n_{10}^{yx}}{\left[(n_{01}^{yx} + n_{11}^{yx}) (n_{10}^{yx} + n_{11}^{yx}) \left(n_{01}^{yx} + \widehat{n_{00}^{yx}} \right) \left(n_{10}^{yx} + \widehat{n_{00}^{yx}} \right) \right]^{1/2}} \quad (1.15)$$

(Chatterjee, 2016) permettant d'évaluer le biais de corrélation. Et donc, l'estimateur donné en (1.13) est adéquat que si $\widehat{\rho^{CD}} = 0$. Dans le cas contraire, $\widehat{n_{00}^{yx}}$ peut être surestimé ($\widehat{\rho^{CD}} > 0$) ou sous-estimé ($\widehat{\rho^{CD}} < 0$).

Que se passe-t-il avec les estimateurs de Lincoln-Petersen de la taille et de la variance si la population sous-étude est hétérogène ? La réponse à cette question n'a pas été donnée de manière explicite dans la littérature. Pour tenter d'y apporter une réponse, on suppose que

l'hétérogénéité de la population se matérialise par l'observation de deux états distinctifs \bar{z} et \underline{z} . On se propose alors d'estimer la taille de la population totale n^{yx} telle que $n^{yx} = n^{yx}(\bar{z}) + n^{yx}(\underline{z})$ où $n^{yx}(\bar{z})$ et $n^{yx}(\underline{z})$ représentent la taille respective des sous-populations correspondant aux états \bar{z} et \underline{z} .

Estimation de la taille de la population totale en présence d'hétérogénéité.

En utilisant l'équation (1.13), la taille estimée $\widehat{n^{yx}(\bar{z})}$ s'écrit :

$$\begin{aligned}\widehat{n^{yx}(\bar{z})} &= n_{01}^{yx}(\bar{z}) + n_{10}^{yx}(\bar{z}) + n_{11}^{yx}(\bar{z}) + \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} \\ &= \frac{1}{n_{11}^{yx}(\bar{z})} [n_{01}^{yx}(\bar{z}) + n_{11}^{yx}(\bar{z})] [n_{10}^{yx}(\bar{z}) + n_{11}^{yx}(\bar{z})].\end{aligned}\quad (1.16)$$

et la taille estimée $\widehat{n^{yx}(\underline{z})}$

$$\begin{aligned}\widehat{n^{yx}(\underline{z})} &= n_{01}^{yx}(\underline{z}) + n_{10}^{yx}(\underline{z}) + n_{11}^{yx}(\underline{z}) + \frac{n_{01}^{yx}(\underline{z}) n_{10}^{yx}(\underline{z})}{n_{11}^{yx}(\underline{z})} \\ &= \frac{1}{n_{11}^{yx}(\underline{z})} [n_{01}^{yx}(\underline{z}) + n_{11}^{yx}(\underline{z})] [n_{10}^{yx}(\underline{z}) + n_{11}^{yx}(\underline{z})].\end{aligned}\quad (1.17)$$

Ainsi, la taille estimée de la population totale $\widehat{n^{yx}}$ tenant compte de l'hétérogénéité est donnée par :

$$\widehat{n^{yx}} = n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx} + \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} + \frac{n_{01}^{yx}(\underline{z}) n_{10}^{yx}(\underline{z})}{n_{11}^{yx}(\underline{z})}.\quad (1.18)$$

Par suite, la non prise en compte de l'hétérogénéité est à l'origine d'un biais $b(n^{yx})$ donné par (équation 1.13 comparée à l'équation 1.18) :

$$b(n^{yx}) = \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} - \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(\underline{z}) n_{10}^{yx}(\underline{z})}{n_{11}^{yx}(\underline{z})}.\quad (1.19)$$

En admettant que $n^{yx} = n^{yx}(\bar{z}) + n^{yx}(\underline{z})$, le biais $b(n^{yx})$ se ré-écrit

$$b(n^{yx}) = \frac{(n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(\underline{z})) (n_{10}^{yx}(\bar{z}) + n_{10}^{yx}(\underline{z}))}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(\underline{z})} - \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(\underline{z}) n_{10}^{yx}(\underline{z})}{n_{11}^{yx}(\underline{z})}.$$

Ce qui permet de ré-écrire l'expression du biais sous la forme (Annexe 1.A.3)

$$b(n^{yx}) = -\frac{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \left(\frac{n_{01}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(z)}{n_{11}^{yx}(z)} \right) \left(\frac{n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \right).$$

Par conséquent, le biais dû à la non prise en compte de l'hétérogénéité disparaîtra que si $n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) = n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})$ et/ou $n_{10}^{yx}(\bar{z}) n_{11}^{yx}(z) = n_{10}^{yx}(z) n_{11}^{yx}(\bar{z})$ ou de manière équivalente si :

$$\frac{n_{11}^{yx}(z)}{n_{01}^{yx}(z)} = \frac{n_{11}^{yx}(\bar{z})}{n_{01}^{yx}(\bar{z})} \quad \text{et/ou} \quad \frac{n_{11}^{yx}(z)}{n_{10}^{yx}(z)} = \frac{n_{11}^{yx}(\bar{z})}{n_{10}^{yx}(\bar{z})}. \quad (1.20)$$

Autrement dit, le biais dû à la non prise en compte de l'hétérogénéité sera nul que si le rapport, pour l'état z , entre le nombre d'individus observés sur les deux listes au nombre d'individus observés seulement sur la liste deux est identique, pour l'état \bar{z} , au nombre d'individus observés sur les deux listes rapporté au nombre d'individus observés seulement sur la deuxième liste.

En présence d'hétérogénéité, la variance de l'estimateur de la population totale pour chaque état, en référence à l'équation (1.14) a pour expression respective :

$$\begin{aligned} V(\widehat{n^{yx}(\bar{z})}) &= \left(\frac{n_{01}^{yx}(\bar{z}) + n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} \right) \left(\frac{n_{10}^{yx}(\bar{z}) + n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} \right) \left(\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} \right) \\ V(\widehat{n^{yx}(z)}) &= \left(\frac{n_{01}^{yx}(z) + n_{11}^{yx}(z)}{n_{11}^{yx}(z)} \right) \left(\frac{n_{10}^{yx}(z) + n_{11}^{yx}(z)}{n_{11}^{yx}(z)} \right) \left(\frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \right) \end{aligned} \quad (1.21)$$

Variance de la taille estimée de la population totale en présence d'hétérogénéité. En présence d'hétérogénéité, la variance estimée de la taille de la population totale est donnée par :

$$\begin{aligned} V(\widehat{n^{yx}}) &= V(\widehat{n^{yx}(\bar{z})} + \widehat{n^{yx}(z)}) \\ &= V(\widehat{n^{yx}(\bar{z})}) + V(\widehat{n^{yx}(z)}) + 2Cov(\widehat{n^{yx}(\bar{z})}, \widehat{n^{yx}(z)}) \\ &= V(\widehat{n_{00}^{yx}(\bar{z})}) + V(\widehat{n_{00}^{yx}(z)}) + 2Cov(\widehat{n_{00}^{yx}(\bar{z})}, \widehat{n_{00}^{yx}(z)}) \end{aligned}$$

car, pour $\omega \in \{01, 10, 11\}$, les tailles $n_{\omega}^{yx}(\bar{z})$ et $n^{yx}(z)$ sont supposées constantes. Ainsi, la covariance $Cov\left(\widehat{n_{00}^{yx}(\bar{z})}, \widehat{n_{00}^{yx}(z)}\right)$ est donnée par :

$$Cov\left(\widehat{n_{00}^{yx}(\bar{z})}, \widehat{n_{00}^{yx}(z)}\right) = -n^{yx} \widehat{p_{00}^{yx}(\bar{z})} \widehat{p_{00}^{yx}(z)}$$

où, en référence aux données du tableau 1.2,

$$\begin{aligned} \widehat{p_{00}^{yx}(\bar{z})} &= (1 - r_1^{yx}(\bar{z}))(1 - r_2^{yx}(\bar{z})) \\ &= \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{[n_{11}^{yx}(\bar{z}) + n_{01}^{yx}(\bar{z})][n_{11}^{yx}(\bar{z}) + n_{10}^{yx}(\bar{z})]} \\ \widehat{p_{00}^{yx}(z)} &= (1 - r_1^{yx}(z))(1 - r_2^{yx}(z)) \\ &= \frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{[n_{11}^{yx}(z) + n_{01}^{yx}(z)][n_{11}^{yx}(z) + n_{10}^{yx}(z)]}. \end{aligned}$$

Par suite,

$$\begin{aligned} Cov\left(\widehat{n_{00}^{yx}(\bar{z})}, \widehat{n_{00}^{yx}(z)}\right) &= -[n^{yx}(\bar{z}) + n^{yx}(z)] \widehat{p_{00}^{yx}(\bar{z})} \widehat{p_{00}^{yx}(z)} \\ &= -n^{yx}(\bar{z}) \widehat{p_{00}^{yx}(\bar{z})} \widehat{p_{00}^{yx}(z)} - n^{yx}(z) \widehat{p_{00}^{yx}(\bar{z})} \widehat{p_{00}^{yx}(z)} \\ &= -\widehat{n_{00}^{yx}(\bar{z})} \widehat{p_{00}^{yx}(z)} - \widehat{n_{00}^{yx}(z)} \widehat{p_{00}^{yx}(\bar{z})} \\ &= -\left(\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})}\right) \widehat{p_{00}^{yx}(z)} - \left(\frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)}\right) \widehat{p_{00}^{yx}(\bar{z})}. \\ &= -\left(\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})}\right) \left\{ \frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{[n_{11}^{yx}(z) + n_{01}^{yx}(z)][n_{11}^{yx}(z) + n_{10}^{yx}(z)]} \right\} \\ &\quad - \left(\frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)}\right) \left\{ \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{[n_{11}^{yx}(\bar{z}) + n_{01}^{yx}(\bar{z})][n_{11}^{yx}(\bar{z}) + n_{10}^{yx}(\bar{z})]} \right\} \\ &= -\left(\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})}\right) \left\{ \frac{\frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)}}{[n_{11}^{yx}(z) + n_{01}^{yx}(z)][n_{11}^{yx}(z) + n_{10}^{yx}(z)]}}{\frac{n_{11}^{yx}(z)}{n_{11}^{yx}(z)}} \right\} \\ &\quad - \left(\frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)}\right) \left\{ \frac{\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})}}{[n_{11}^{yx}(\bar{z}) + n_{01}^{yx}(\bar{z})][n_{11}^{yx}(\bar{z}) + n_{10}^{yx}(\bar{z})]} \right\}. \end{aligned}$$

Ainsi,

$$\begin{aligned} \text{Cov} \left(\widehat{n_{00}^{yx}(\bar{z})}, \widehat{n_{00}^{yx}(z)} \right) &= - \left(\frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z}) n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)} \right) \left(\frac{1}{\widehat{n^{yx}(z)}} + \frac{1}{\widehat{n^{yx}(\bar{z})}} \right) \\ &= -\widehat{n_{00}^{yx}(\bar{z})} \widehat{n_{00}^{yx}(z)} \left[\frac{\widehat{n^{yx}(\bar{z})} + \widehat{n^{yx}(z)}}{\widehat{n^{yx}(\bar{z})} \widehat{n^{yx}(z)}} \right] < 0. \end{aligned}$$

Par conséquent, la prise en compte de l'hétérogénéité fait augmenter la variance totale c'est-à-dire :

$$V(\widehat{n^{yx}}) < V(\widehat{n^{yx}(\bar{z})}) + V(\widehat{n^{yx}(z)}).$$

Estimateur de Chapman. Une modification de l'estimateur de Lincoln-Petersen a été proposée par Chapman en 1951 pour pallier au problème de non chevauchement potentiel entre les deux listes. Cet estimateur (Brittain et Bohning, 2009), encore appelé estimateur modifié de Lincoln-Petersen, est tel que

$$\widehat{n_{00}^{yx}} = \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx} + 1}.$$

En conséquence,

$$\widehat{n^{yx}} = \frac{1}{n_{11}^{yx} + 1} (n_{01}^{yx} + n_{11}^{yx} + 1) (n_{10}^{yx} + n_{11}^{yx} + 1). \quad (1.22)$$

La variance asymptotique de cet estimateur, déterminée par Seber en 1970, a pour expression :

$$V(\widehat{n^{yx}}) = \frac{(n_{01}^{yx} + n_{11}^{yx} + 1) (n_{10}^{yx} + n_{11}^{yx} + 1) (n_{01}^{yx} n_{10}^{yx})}{(n_{11}^{yx} + 1)^2 (n_{11}^{yx} + 2)}. \quad (1.23)$$

Estimateur de la limite inférieure de Chao. Chao (1989) a proposé un estimateur qui se libère de l'hypothèse de l'indépendance des deux sources. L'expression de l'estimateur de n_{00}^{yx} dans ces conditions est donnée par :

$$\widehat{n_{00}^{yx}} = \frac{1}{4} \frac{(n_{10}^{yx} + n_{01}^{yx})^2}{n_{11}^{yx}}$$

de sorte que

$$\widehat{n^{yx}} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} + \frac{1}{4} \frac{(n_{10}^{yx} + n_{01}^{yx})^2}{n_{11}^{yx}}. \quad (1.24)$$

L'expression de la variance de l'estimateur proposé par Chao est similaire à celle de l'estimateur de Zelterman donnée à l'équation (1.26) au paragraphe suivant.

Estimateur de Zelterman. L'estimateur de Zelterman est basé sur l'estimateur de la taille d'une population de Horvitz–Thompson, dérivé à partir d'une distribution mixte de Poisson tronquée en zéro. Pour deux sources, Brittain et Bohning (2009) ont dérivé un estimateur du maximum de vraisemblance de n_{00}^{yx} et sa variance, basé sur une distribution binomiale mixte tronquée en zéro. Cet estimateur a pour expression :

$$\widehat{n}_{00}^{yx} = \left(\frac{n_{10}^{yx} + n_{01}^{yx}}{n_{10}^{yx} + n_{01}^{yx} + 2n_{11}^{yx}} \right)^2.$$

Après,

$$\widehat{n}^{yx} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} + \left(\frac{n_{10}^{yx} + n_{01}^{yx}}{n_{10}^{yx} + n_{01}^{yx} + 2n_{11}^{yx}} \right)^2. \quad (1.25)$$

L'expression de la variance de l'estimateur de Zelterman est donnée par :

$$V(\widehat{n}^{yx}) = \frac{1}{4} \left[\frac{(n_{01}^{yx} + n_{10}^{yx})^2}{n_{11}^{yx}} \right] \left(\frac{n_{01}^{yx} + n_{10}^{yx}}{2n_{11}^{yx}} + 1 \right)^2. \quad (1.26)$$

Brittain et Bohning (2009) ont caractérisé et comparé ces estimateurs. Ils ont établi sous certaines conditions l'équivalence entre les estimateurs de Lincoln-Petersen, Chapman et Chao. Ils ont proposé deux autres estimateurs : un estimateur du maximum de vraisemblance et un estimateur de McKendricks déterminé par la méthode des moments. Les simulations effectuées ont mis en avant la supériorité de l'estimateur développé par Chao (1989, 1987). Cette chercheuse a aussi élaboré une mesure de dépendance, le coefficient de covariation, entre les sources 1 et 2 permettant d'évaluer l'hypothèse d'indépendance. Elle a utilisé cette mesure pour évaluer théoriquement et par simulation les estimateurs de Lincoln-Petersen et de Chapman (Chao et Chiang, 2008).

Malgré ses limites, l'estimateur de Lincoln-Petersen continue d'être utilisé et fournit dans de nombreux cas des estimations fiables (El Allaki et Vallières, 2015; Xu et He, 2014; Khan et Uddin, 2004). Pour estimer les paramètres de notre modèle de régression, nous adoptons l'estimateur de Lincoln-Petersen en supposant que $n_{11}^{yx} > 0$. Sa variance sera

approximée par la méthode delta en relaxant l'hypothèse de constance de $n_{11}^{yx} + n_{10}^{yx}$ et $n_{11}^{yx} + n_{01}^{yx}$ et en maintenant l'hypothèse d'indépendance des deux sources (Chao et Tsay, 1998).

1.4.2.3 Les estimateurs des paramètres du modèle

Les estimateurs ponctuels. Les estimateurs de \bar{q} , \underline{q} et \hat{q} définis respectivement dans les équations (1.6), (1.7), (1.8) et les estimateurs de r_1^{yx} et r_2^{yx} du modèle sont donnés par (Lavallée et Rivest, 2013; Rivest, 2013) :

$$\hat{q} = \frac{\widehat{n^{11}}}{\widehat{n^{11}} + \widehat{n^{01}}}, \quad (1.27)$$

$$\underline{q} = \frac{\widehat{n^{10}}}{\widehat{n^{10}} + \widehat{n^{00}}}, \quad (1.28)$$

$$\hat{q} = \frac{\widehat{n^{11}} + \widehat{n^{01}}}{\widehat{N}}, \quad (1.29)$$

et, pour $yx \in \Omega = \{00, 01, 10, 11\}$,

$$\widehat{r_1^{yx}} = \frac{n_{11}^{yx} + n_{10}^{yx}}{\widehat{n^{yx}}}, \quad \widehat{r_2^{yx}} = \frac{n_{11}^{yx} + n_{01}^{yx}}{\widehat{n^{yx}}} \quad (1.30)$$

où l'expression de $\widehat{n^{yx}}$ est donnée à l'équation (1.13) et celle de $\widehat{n_{00}^{yx}}$ à l'équation (1.12).

Etant donné l'expression de l'estimateur de n_{00}^{yx} (équation 1.12), les estimateurs de \bar{q}

et \underline{q} sont alors respectivement donnés par :

$$\begin{aligned}\widehat{q} &= \frac{\widehat{n^{11}}}{\widehat{n^{11}} + \widehat{n^{01}}}, \\ &= \frac{n_{01}^{11} + n_{10}^{11} + n_{11}^{11} + \frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}}}{n_{01}^{11} + n_{10}^{11} + n_{11}^{11} + \frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}} + n_{01}^{01} + n_{10}^{01} + n_{11}^{01} + \frac{n_{01}^{01}n_{10}^{01}}{n_{11}^{01}}}, \\ &= \frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + (n_{11}^{11}/n_{11}^{01})(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}\end{aligned}\quad (1.31)$$

$$\begin{aligned}\widehat{\underline{q}} &= \frac{\widehat{n^{10}}}{\widehat{n^{10}} + \widehat{n^{00}}}, \\ &= \frac{n_{01}^{10} + n_{10}^{10} + n_{11}^{10} + \frac{n_{01}^{10}n_{10}^{10}}{n_{11}^{10}}}{n_{01}^{10} + n_{10}^{10} + n_{11}^{10} + \frac{n_{01}^{10}n_{10}^{10}}{n_{11}^{10}} + n_{01}^{00} + n_{10}^{00} + n_{11}^{00} + \frac{n_{01}^{00}n_{10}^{00}}{n_{11}^{00}}}, \\ &= \frac{(n_{01}^{10} + n_{11}^{10})(n_{10}^{10} + n_{11}^{10})}{(n_{01}^{10} + n_{11}^{10})(n_{10}^{10} + n_{11}^{10}) + (n_{11}^{10}/n_{11}^{00})(n_{01}^{00} + n_{11}^{00})(n_{10}^{00} + n_{11}^{00})}.\end{aligned}\quad (1.32)$$

Variance des estimateurs $\widehat{n^{yx}}$, \widehat{q} et $\widehat{\underline{q}}$.

Variance asymptotique de $\widehat{n^{yx}}$. En supposant que les n_{ω}^{yx} sont des binomiales et que $n_{11}^{yx} + n_{10}^{yx}$ et $n_{11}^{yx} + n_{01}^{yx}$ ne sont pas constants, l'expression de la variance de $\widehat{n^{yx}}$, approximée par la méthode delta (Annexe : 1.A.4) est équivalente à ⁸ :

$$\begin{aligned}V(\widehat{n^{yx}}) &\simeq n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right)^2 + \frac{1}{n_{11}^{yx}} \left(\widehat{n^{y0}} - 1\right)^2 \\ &\quad - \left\{ \frac{\left(\widehat{n^{y0}} - 1\right)^2}{\widehat{n^{yx}}} + \frac{1}{n_{11}^{yx}} \left[n_{01}^{yx} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) + n_{10}^{yx} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right) \right] \right\}\end{aligned}\quad (1.33)$$

Variance de \widehat{q} et $\widehat{\underline{q}}$. L'évaluation de la précision de ces estimateurs et la réalisation d'éventuels tests y relatifs impliquent la connaissance de leur variance. On se propose dans ce paragraphe d'estimer la variance tout au moins asymptotique de ces estimateurs.

8. Cet estimateur $V(\widehat{n^{yx}})$ de la variance peut être instable, car le nombre d'individus (n_{11}) apparaissant plusieurs fois figure au dénominateur de plusieurs quantités. Cette remarque s'applique tout aussi bien aux différentes expressions de $V(\widehat{n^{yx}})$ établies dans les différentes publications où l'expression de $V(\widehat{n^{yx}})$ est toujours réduite par une fonction $f(n_{11}^{yx})$. En effet, $f(n_{11}^{yx}) = (n_{11}^{yx})^3$ (Sekar et Deming, 1949) ; $f(n_{11}^{yx}) = (n_{11}^{yx} + 1)^2 (n_{11}^{yx} + 2)$ (Chapman, 1951 ; Chao, 1989).

L'expression de l'estimateur \widehat{q} (équation 1.31) s'écrit :

$$\widehat{q} = \frac{\widehat{n}^{11}}{\widehat{n}^{11} + \widehat{n}^{01}}.$$

Cette expression étant non linéaire, la variance de \widehat{q} est approximée au voisinage de

$$\mu = \left(E\left(\widehat{n}^{11}\right), E\left(\widehat{n}^{01}\right) \right)$$

par la *Méthode Delta* (Mittelhammer, 2013). Par suite,

$$\begin{aligned} V(\widehat{q}) &= V\left(\frac{\widehat{n}^{11}}{\widehat{n}^{11} + \widehat{n}^{01}}\right) \\ &\simeq \frac{\left[E\left(\widehat{n}^{01}\right)\right]^2}{\left[E\left(\widehat{n}^{11}\right) + E\left(\widehat{n}^{01}\right)\right]^4} V\left(\widehat{n}^{11}\right) + \frac{\left[E\left(\widehat{n}^{11}\right)\right]^2}{\left[E\left(\widehat{n}^{11}\right) + E\left(\widehat{n}^{01}\right)\right]^4} V\left(\widehat{n}^{01}\right) \end{aligned} \quad (1.34)$$

où $Cov\left(\widehat{n}_{00}^{01}, \widehat{n}_{00}^{11}\right) = 0$ en accord avec l'hypothèse *d'indépendance* des observations faites dans chacune des deux listes. Les estimateurs \widehat{n}^{yx} et $V\left(\widehat{n}^{yx}\right)$ sont donnés respectivement aux équations (1.13) et (1.33).

De manière similaire, une approximation de la variance de \widehat{q} au voisinage de

$$\mu = \left(E\left(\widehat{n}^{10}\right), E\left(\widehat{n}^{00}\right) \right)$$

est telle que

$$\begin{aligned} V(\widehat{q}) &= V\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right) \\ &\simeq \frac{\left[E\left(\widehat{n}^{00}\right)\right]^2}{\left[E\left(\widehat{n}^{10}\right) + E\left(\widehat{n}^{00}\right)\right]^4} V\left(\widehat{n}^{10}\right) + \frac{\left[E\left(\widehat{n}^{10}\right)\right]^2}{\left[E\left(\widehat{n}^{10}\right) + E\left(\widehat{n}^{00}\right)\right]^4} V\left(\widehat{n}^{00}\right). \end{aligned} \quad (1.35)$$

Un regard sur l'estimateur naïf relatif à \widehat{q} et \widehat{q} . On cherche à vérifier si les observations faites dans les deux listes ne seraient pas suffisantes pour estimer \bar{q} et q . Autrement dit, si on dispose des observations n_{ω}^{yx} , $\omega \in \{01, 10, 11\}$ et $yx \in \{00, 01, 10, 11\}$, ne peut-on

pas estimer de manière fiable \bar{q} et \underline{q} en absence de l'estimation de n_{00}^{yx} ?

Si l'on ne tient pas compte des problèmes d'échantillonnage, les probabilités conditionnelles \bar{q} et \underline{q} sont obtenues par les estimateurs 'naïfs',⁹ suivants :

$$\tilde{q} = \frac{o^{11}}{o^{11} + o^{01}} \quad \text{et} \quad \underline{\tilde{q}} = \frac{o^{10}}{o^{10} + o^{00}} \quad (1.36)$$

où l'on note $o^{yx} = n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}$ le total des individus de caractéristiques yx observés.

Calcul du biais par rapport aux estimateurs de \bar{q} et \underline{q} . En notant $o^{yx} = n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}$, l'estimateur de \bar{q} donné à l'équation (1.31) peut se réécrire sous la forme :

$$\widehat{\bar{q}} = \frac{o^{11} + \widehat{n}_{00}^{11}}{o^{11} + \widehat{n}_{00}^{11} + o^{01} + \widehat{n}_{00}^{01}}.$$

L'estimateur naïf \tilde{q} est donc à l'origine d'un biais $b_{\bar{q}}$ donné par :

$$\begin{aligned} b_{\bar{q}} &= \tilde{q} - \widehat{\bar{q}} \\ &= \frac{o^{11}}{o^{11} + o^{01}} - \frac{o^{11} + \widehat{n}_{00}^{11}}{o^{11} + \widehat{n}_{00}^{11} + o^{01} + \widehat{n}_{00}^{01}} \\ &= \left(\frac{o^{11} o^{01}}{o^{11} + o^{01}} \right) \left[\frac{\left(\widehat{n}_{00}^{01} / o^{01} \right) - \left(\widehat{n}_{00}^{11} / o^{11} \right)}{o^{11} + \widehat{n}_{00}^{11} + o^{01} + \widehat{n}_{00}^{01}} \right]. \end{aligned} \quad (1.37)$$

De manière similaire, l'estimateur naïf $\underline{\tilde{q}}$ est à l'origine d'un biais $b_{\underline{q}}$ défini par :

$$\begin{aligned} b_{\underline{q}} &= \underline{\tilde{q}} - \widehat{\underline{q}} \\ &= \frac{o^{10}}{o^{10} + o^{00}} - \frac{o^{10} + \frac{n_{01}^{10} n_{10}^{10}}{n_{11}^{10}}}{o^{10} + \widehat{n}_{00}^{10} + o^{00} + \widehat{n}_{00}^{00}} \\ &= \left(\frac{o^{00} o^{10}}{o^{10} + o^{00}} \right) \left[\frac{\left(\widehat{n}_{00}^{00} / o^{00} \right) - \left(\widehat{n}_{00}^{10} / o^{10} \right)}{o^{10} + \widehat{n}_{00}^{10} + o^{00} + \widehat{n}_{00}^{00}} \right]. \end{aligned} \quad (1.38)$$

Les formules (1.37) et (1.38) et la formule de l'estimateur de n_{00}^{yx} donnée en (1.12) permettent d'établir que le biais des estimateurs naïfs \tilde{q} et $\underline{\tilde{q}}$ est nul si et seulement si

9. Un estimateur naïf est un estimateur défini sur la base des seules observations, sans prendre en compte les problèmes d'échantillonnage.

respectivement on a :

$$\left(\frac{o^{11}}{o^{01}}\right) = \frac{\widehat{n_{00}^{11}}}{\widehat{n_{00}^{01}}} = \left(\frac{n_{01}^{11}n_{10}^{11}}{n_{01}^{01}n_{10}^{01}}\right) \left(\frac{n_{11}^{11}}{n_{11}^{01}}\right)^{-1}, \quad (1.39)$$

$$\left(\frac{o^{10}}{o^{00}}\right) = \frac{\widehat{n_{00}^{10}}}{\widehat{n_{00}^{00}}} = \left(\frac{n_{01}^{10}n_{10}^{10}}{n_{01}^{00}n_{10}^{00}}\right) \left(\frac{n_{11}^{10}}{n_{11}^{00}}\right)^{-1}. \quad (1.40)$$

Ces deux conditions peuvent se ré-écrire :

$$\frac{\widehat{n_{00}^{11}}}{\widehat{n_{11}^{11}}} = \frac{n_{01}^{11}n_{10}^{11}}{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})} = \frac{n_{01}^{01}n_{10}^{01}}{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})} = \frac{\widehat{n_{00}^{01}}}{\widehat{n_{01}^{01}}}, \quad (1.41)$$

$$\frac{\widehat{n_{00}^{10}}}{\widehat{n_{10}^{10}}} = \frac{n_{01}^{10}n_{10}^{10}}{(n_{01}^{10} + n_{11}^{10})(n_{10}^{10} + n_{11}^{10})} = \frac{n_{01}^{00}n_{10}^{00}}{(n_{01}^{00} + n_{11}^{00})(n_{10}^{00} + n_{11}^{00})} = \frac{\widehat{n_{00}^{00}}}{\widehat{n_{00}^{00}}}. \quad (1.42)$$

L'expression (1.41) indique que l'estimateur naïf \widetilde{q} sera non biaisé si et seulement si la proportion des individus non-observés est identique pour les individus de caractéristique $yx = 11$ et ceux de caractéristiques $yx = 01$; autrement dit, tous les individus de caractéristique $x = 1$, ont la même probabilité d'être observés.¹⁰ De même, l'expression (1.42) indique que l'estimateur naïf \widetilde{q} sera non biaisé si et seulement si les individus de caractéristique $x = 0$, ont la même probabilité d'être observés.

Proposition 1.1. *Pour que l'estimateur naïf de \bar{q} soit asymptotiquement sans biais, il faut que la probabilité que les individus ne fassent partie d'aucune des deux listes soit identique pour les individus de caractéristiques $yx = 01$ et ceux de caractéristiques $yx = 11$.*

Démonstration. Annexe : 1.A.5

Le biais $b_{\bar{q}}$ est asymptotiquement nul si et seulement si

$$(1 - r_1^{11})(1 - r_2^{11}) = (1 - r_1^{01})(1 - r_2^{01}). \quad (1.43)$$

En d'autre termes, on retrouve le fait que, pour que l'estimateur naïf de \bar{q} soit (asymptotiquement) sans biais, il faut que la probabilité que les individus ne fassent pas partie d'aucune des deux listes soit identique pour les individus de caractéristiques $yx = 01$ et

¹⁰ Observons que la condition (1.41) n'implique pas forcément que les probabilités de capture r_1^{yx} et r_2^{yx} sont identiques pour tous les individus de caractéristique $x = 1$. Par contre, évidemment, si les probabilités de capture sont identiques, la condition (1.41) est nécessairement vérifiée.

ceux de caractéristiques $yx = 11$ - c'est à dire pour tous les individus de caractéristique $x = 1$. \square

De façon similaire, on peut établir que l'estimateur naïf de \underline{q} est asymptotiquement sans biais si et seulement si

$$(1 - r_1^{10})(1 - r_2^{10}) = (1 - r_1^{00})(1 - r_2^{00}), \quad (1.44)$$

c'est à dire si tous les individus de caractéristique $x = 0$ ont la même probabilité de n'appartenir à aucune des deux listes.

Bien sûr, il n'y a aucune raison de supposer *a priori* que ni la condition (1.43) ni la condition (1.44) soient remplies.

Un test asymptotique de biais nul. Lorsque n^{yx} est suffisamment élevé¹¹ et sous l'hypothèse d'indépendance, on peut tester statistiquement l'hypothèse de biais nul. En d'autres termes, on peut évaluer si l'estimateur naïf est adéquat pour estimer les probabilités conditionnelles \bar{q} et \underline{q} en absence de l'estimation du nombre des individus qui n'ont pas été observés dans les deux listes.

Soit $p_{00}^{11} = (n_{00}^{11}/n^{11})$ et $p_{00}^{01} = (n_{00}^{01}/n^{01})$ respectivement la proportion d'individus de caractéristiques $yx = 11$ et $yx = 01$ absents des deux listes. Les hypothèses statistiques définissant la nullité du biais ($b_{\bar{q}} = 0$) s'écrivent :

$$H_0 : p_{00}^{11} - p_{00}^{01} = 0 \quad (b_{\bar{q}} = 0)$$

$$H_1 : p_{00}^{11} - p_{00}^{01} \neq 0.$$

Pour déterminer la statistique du test, on utilise le fait que $n_{00}^{yx} \sim \mathcal{B}(n^{yx}, \pi_{00}^{yx})$ où $\pi_{00}^{yx} = (1 - r_1^{yx})(1 - r_2^{yx})$, $E(n_{00}^{yx}) = n^{yx}(1 - r_1^{yx})(1 - r_2^{yx})$ et $V(n_{00}^{yx}) = n^{yx}(1 - r_1^{yx})(1 - r_2^{yx}) \times [1 - (1 - r_1^{yx})(1 - r_2^{yx})]$. Disposant des estimateurs $\widehat{n_{\omega}^{yx}}$ et $\widehat{n^{yx}}$ donnés aux équations (1.12) et (1.13), les estimateurs de p_{00}^{11} et de p_{00}^{01} sont donc respectivement donnés par

11. C'est l'hypothèse généralement adoptée. En pratique, on a une bonne approximation quand $n^{yx} \geq 30$ et $\min(n^{yx}\pi_{00}^{yx}, n^{yx}\pi_{00}^{yx}(1 - \pi_{00}^{yx})) \geq 5$. Evidemment, l'approximation est d'autant plus bonne que $n^{yx} \rightarrow +\infty$.

(Annexe : 1.67) :

$$\begin{aligned}\widehat{p}_{00}^{11} &= \frac{\widehat{n}_{00}^{11}}{\widehat{n}^{11}} \\ &= \frac{n_{01}^{11}n_{10}^{11}}{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}\end{aligned}\quad (1.45)$$

$$\begin{aligned}\widehat{p}_{00}^{01} &= \frac{\widehat{n}_{00}^{01}}{\widehat{n}^{01}} \\ &= \frac{n_{01}^{01}n_{10}^{01}}{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}.\end{aligned}\quad (1.46)$$

Par suite, sous l'hypothèse d'indépendance et en utilisant à nouveau la *Méthode Delta* (Mittelhammer, 2013), on obtient l'espérance mathématique

$$E\left(\widehat{p}_{00}^{11}\right) = E\left(\frac{\widehat{n}_{00}^{11}}{\widehat{n}^{11}}\right) \simeq \frac{E\left(\widehat{n}_{00}^{11}\right)}{E\left(\widehat{n}^{11}\right)}$$

et la variance

$$\begin{aligned}V\left(\widehat{p}_{00}^{11}\right) &= V\left(\frac{\widehat{n}_{00}^{11}}{\widehat{n}^{11}}\right) \\ &\simeq \left[E\left(\widehat{n}^{11}\right)\right]^{-2} V\left(\widehat{n}_{00}^{11}\right) + \left\{n_{00}^{11} \left[E\left(\widehat{n}^{11}\right)\right]^{-2}\right\}^2 V\left(\widehat{n}^{11}\right).\end{aligned}$$

De façon similaire,

$$\begin{aligned}E\left(\widehat{p}_{00}^{01}\right) &= E\left(\frac{\widehat{n}_{00}^{01}}{\widehat{n}^{01}}\right) \simeq \frac{E\left(\widehat{n}_{00}^{01}\right)}{E\left(\widehat{n}^{01}\right)} \\ V\left(\widehat{p}_{00}^{01}\right) &= V\left(\frac{\widehat{n}_{00}^{01}}{\widehat{n}^{01}}\right) \\ &\simeq \left[E\left(\widehat{n}^{01}\right)\right]^{-2} V\left(\widehat{n}_{00}^{01}\right) + \left\{n_{00}^{01} \left[E\left(\widehat{n}^{01}\right)\right]^{-2}\right\}^2 V\left(\widehat{n}^{01}\right)\end{aligned}$$

La détermination complète de ces espérances mathématiques et de ces variances passe par le calcul de l'espérance mathématique et de la variance de \widehat{n}_{00}^{yx} . En utilisant la méthode delta, on obtient respectivement pour l'espérance mathématique et la variance de \widehat{n}_{00}^{yx}

(Annexe 1.A.7) :

$$\begin{aligned} E\left(\widehat{n_{00}^{yx}}\right) &= E\left(\frac{n_{10}^{yx}n_{10}^{yx}}{n_{11}^{yx}}\right) \\ &\simeq E\left(n_{00}^{yx}\right) \end{aligned} \quad (1.47)$$

$$\begin{aligned} V\left(\widehat{n_{00}^{yx}}\right) &= V\left(\frac{n_{10}^{yx}n_{10}^{yx}}{n_{11}^{yx}}\right) \\ &\simeq \widehat{n_{00}^{yx}} \left[\frac{1}{n_{11}^{yx}} \left(n_{10}^{yx} + n_{01}^{yx} + \widehat{n_{00}^{yx}} \right) - 3 \frac{\widehat{n_{00}^{yx}}}{n_{11}^{yx}} \right]. \end{aligned} \quad (1.48)$$

Par suite, pour n^{yx} suffisamment grand,

$$\widehat{p_{00}^{yx}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(E\left(\widehat{p_{00}^{01}}\right), V\left(\widehat{p_{00}^{01}}\right)\right).$$

Pour déterminer la variance $V\left(\widehat{p_{00}^{11}} - \widehat{p_{00}^{01}}\right)$, on pose $p_{00}^{11} = p_{00}^{01} = \bar{p}$ sous l'hypothèse H_0 . Un estimateur de \bar{p} et sa variance peuvent être calculés de la manière suivante (Annexe 1.A.7.2, équations 1.68 et 1.69) :

$$\begin{aligned} \widehat{\bar{p}} &= \frac{\widehat{n_{11}^{11}}}{\widehat{n_{11}^{11}} + \widehat{n_{01}^{01}}} \widehat{p_{00}^{11}} + \frac{\widehat{n_{01}^{01}}}{\widehat{n_{11}^{11}} + \widehat{n_{01}^{01}}} \widehat{p_{00}^{01}} \\ &= \frac{n_{11}^{01} (n_{01}^{11} n_{10}^{11}) + n_{11}^{11} (n_{01}^{01} n_{10}^{01})}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})}, \end{aligned} \quad (1.49)$$

$$\begin{aligned} V\left(\widehat{\bar{p}}\right) &\equiv V\left(\widehat{p_{00}^{11}} - \widehat{p_{00}^{01}}\right) = \widehat{\bar{p}} \left(1 - \widehat{\bar{p}}\right) \left[\left(\widehat{n_{11}^{11}}\right)^{-1} + \left(\widehat{n_{01}^{01}}\right)^{-1} \right] \\ &= \frac{(n_{11}^{01} n_{01}^{11} n_{10}^{11} + n_{11}^{11} n_{01}^{01} n_{10}^{01})^2 [(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})]}{[n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})]^2}. \end{aligned} \quad (1.50)$$

L'expression de $\widehat{p_{00}^{11}} - \widehat{p_{00}^{01}}$ peut s'écrire

$$\begin{aligned} \widehat{p_{00}^{11}} - \widehat{p_{00}^{01}} &= \frac{n_{01}^{11} n_{10}^{11}}{(n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11})} - \frac{n_{01}^{01} n_{10}^{01}}{(n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \\ &= \frac{n_{01}^{11} n_{10}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01}) - n_{01}^{01} n_{10}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11})}{(n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \\ &= \frac{n_{01}^{11} n_{10}^{11} [n_{01}^{01} (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})] - n_{01}^{01} n_{10}^{01} [n_{01}^{11} (n_{01}^{11} + n_{10}^{11} + n_{11}^{11})]}{(n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \end{aligned} \quad (1.51)$$

de sorte que ¹²

$$Z = \frac{\widehat{p}_{00}^{11} - \widehat{p}_{00}^{01}}{\sqrt{V(\widehat{p}_{00}^{11} - \widehat{p}_{00}^{01})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Les observations fournissent de l'évidence que l'estimateur naïf est adéquat pour estimer les probabilités conditionnelles à un risque α ($0 < \alpha < 1$) que si :

$$P\left(\left|\frac{\widehat{p}_{00}^{11} - \widehat{p}_{00}^{01}}{\sqrt{V(\widehat{p}_{00}^{11} - \widehat{p}_{00}^{01})}}\right| \leq z_{\alpha/2}\right) = 1 - \alpha$$

où $z_{\alpha/2}$ est le quantile d'ordre α tel que pour $Z \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$, $P(|Z| \leq z_{\alpha/2}) = 1 - \alpha$.

Variance des estimateurs naïfs \widetilde{q} et \underline{q} . En se référant au processus sous-jacent générateur des données, on peut calculer la variance de $o^{yx} = n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}$ et en déduire les variances de \widetilde{q} et \underline{q} . L'objectif final est de comparer ces variances à celles de \widehat{q} et de \underline{q} .

Pour déterminer l'espérance mathématique et la variance de o^{yx} , il nous faut déterminer la distribution conjointe de $(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})$ supposant celle de n_{00}^{yx} fixée. En réalité, l'estimateur naïf suppose implicitement la non observation de n_{00}^{yx} bien que celui-ci existe. Sous ces considérations, la distribution conjointe de $(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})$ étant donné celle de n_{00}^{yx} est donnée à l'annexe (1.A.6.1). D'après le processus générateur des données déterminé à l'annexe (1.A.6.1), les $\widetilde{n}_{\omega}^{yx}$ sont des variables binomiales. De manière précise,

$$\begin{aligned}\widetilde{n}_{01}^{yx} &\sim \mathcal{B}(\widetilde{n}^{yx}, \widetilde{p}_{01}^{yx}) \\ \widetilde{n}_{10}^{yx} &\sim \mathcal{B}(\widetilde{n}^{yx}, \widetilde{p}_{10}^{yx}) \\ \widetilde{n}_{11}^{yx} &\sim \mathcal{B}(\widetilde{n}^{yx}, \widetilde{p}_{11}^{yx})\end{aligned}\tag{1.52}$$

12. Dans le cas de petits échantillons, le théorème de la limite centrale n'est pas applicable. Différents tests sont alors proposés. Deux des tests les plus utilisés sont le test *khi-deux* de Pearson et le test exact de Fisher. En réalité, ces deux tests sont utilisés pour tester l'indépendance entre une variable X (en lignes) et une variable Y (en colonnes) rangées dans un tableau de contingence de r lignes et c colonnes. Dénommés aussi modèle à essais comparatifs, de double dichotomie ou modèle d'homogénéité, basés sur une double binomiale, ils sont aussi utilisés pour comparer deux proportions (Sulewski, 2019; Kroll, 1989; Yates, 1984). Dans le cas du test *khi-deux* de Pearson, une correction de continuité est réalisée pour l'adapter au cas de petits échantillons quand au moins une des cellules du tableau de contingence a un effectif inférieur à 5 sous la contrainte que les effectifs partiels théoriques soient tous supérieurs à 3 (Yates, 1984).

où $\tilde{n}^{yx} = n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}$ et $\tilde{p}_{\omega}^{yx} = p_{\omega}^{yx} / (1 - p_{00}^{yx})$.

1. *Espérance mathématique et variance de o^{yx}* . (Annexe 1.A.6.1)

$$\begin{aligned} E(o^{yx}) &= E(n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}) \\ &= n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx} \end{aligned} \quad (1.53)$$

En supposant les variables $n_{\omega}^{yx} / n_{00}^{yx}$, $\omega \in \{01, 10, 11\}$ indépendantes, la variance $V(o^{yx})$ (Annexe 1.A.6.2) est donnée par :

$$\begin{aligned} V(o^{yx}) &= V(n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}) \\ &= 2n_{11}^{yx} \left\{ 1 - \left[\left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}} \right) \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right) \right]^{-1} \right\}. \end{aligned} \quad (1.54)$$

2. *Variance des estimateurs de \tilde{q} et \underline{q}* . En suivant la démarche utilisée pour établir les variances établies aux équations (1.34) et (1.35), les variances de \tilde{q} et \underline{q} sont respectivement approximées par :

$$\begin{aligned} V(\tilde{q}) &= V\left(\frac{o^{11}}{o^{11} + o^{01}}\right) \\ &\simeq \frac{[E(o^{01})]^2}{[E(o^{11}) + E(o^{01})]^4} V(o^{11}) + \frac{[E(o^{11})]^2}{[E(o^{11}) + E(o^{01})]^4} V(o^{01}) \end{aligned} \quad (1.55)$$

et

$$\begin{aligned} V(\underline{q}) &= V\left(\frac{o^{10}}{o^{10} + o^{00}}\right) \\ &\simeq \frac{[E(o^{00})]^2}{[E(o^{10}) + E(o^{00})]^4} V(o^{10}) + \frac{[E(o^{10})]^2}{[E(o^{10}) + E(o^{00})]^4} V(o^{00}) \end{aligned} \quad (1.56)$$

où $Cov(o^{11}, o^{01}) = Cov(o^{10}, o^{00}) = 0$.

Comparaison de $V(\hat{q})$ et de $V(\tilde{q})$. En comparant dans un premier temps deux à deux les coefficients des deux variances, il s'avère que les coefficients de $V(\tilde{q})$ sont

inférieurs à ceux de $V(\widehat{q})$ que si et seulement si (Annexe : 1.A.6.3)

$$\frac{o^{01}}{o^{11}} < \frac{E\left(\widehat{n}_{00}^{01}\right)}{E\left(\widehat{n}_{00}^{11}\right)}. \quad (1.57)$$

En d'autres termes, les coefficients de $V(\widetilde{q})$ sont supérieurs à ceux de $V(\widehat{q})$ si le rapport entre le nombre moyen des individus observés de caractéristique $yx = 01$ et ceux de caractéristique $yx = 11$ est inférieur au rapport entre le nombre moyen des individus non observés de caractéristique $yx = 01$ et ceux de caractéristique $yx = 11$.

Dans un deuxième temps, la comparaison des variances $V(\widehat{n}^{yx})$ et $V(o^{yx})$ (Annexe 1.A.6.3) déterminées aux équations (1.33) et (1.54) permet de conclure que $V(o^{yx}) < V(\widehat{n}^{yx})$. Il s'ensuit que $V(\widetilde{q}) < V(\widehat{q})$.

Par conséquent, si les conditions (1.43) et (1.57) sont vérifiées, mieux vaut utiliser l'estimateur naïf car il se révèle plus précis.

1.4.2.4 Le modèle estimé

L'objectif était d'estimer le modèle établi à l'équation (1.4), re-écriture du modèle $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$ (équation 1.1) à partir de la règle statistique définie au système (1.3). Disposant des estimateurs \widehat{q} et \underline{q} et leur variance après avoir réévalué la variance de \widehat{n}^{yx} en relaxant l'hypothèse de constance de $n_{10}^{yx} + n_{11}^{yx}$ et $n_{01}^{yx} + n_{11}^{yx}$, les paramètres β_0 et β_1 ont été estimés. Les estimateurs obtenus à partir des méthodes de capture recapture (*EMCR*) sont les suivants (Φ étant la fonction de répartition connue de ε_i) :

$$\begin{aligned} \widehat{\beta}_0 &\equiv \Phi^{-1}(\widehat{q}) = \Phi^{-1}\left(P(y_i = 1/x = 0)\right) \\ &= \Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right) \end{aligned}$$

et

$$\begin{aligned}\widehat{\beta}_1 &\equiv \Phi^{-1}(\widehat{\widehat{q}}) - \Phi^{-1}(\widehat{q}) = \Phi^{-1}\left(P(y_i = \widehat{1/x} = 1)\right) - \Phi^{-1}\left(P(y_i = \widehat{1/x} = 0)\right) \\ &= \Phi^{-1}\left(\frac{\widehat{n}^{11}}{\widehat{n}^{11} + \widehat{n}^{01}}\right) - \Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right)\end{aligned}$$

Le modèle estimé s'écrit (voir équation 1.4) :

$$g(\widehat{p}_i) = \Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right) + \left[\Phi^{-1}\left(\frac{\widehat{n}^{11}}{\widehat{n}^{11} + \widehat{n}^{01}}\right) - \Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right)\right] x_i.$$

Les estimateurs $\widehat{\beta}_0$ et $\widehat{\beta}_1$ sont sans biais. En effet,

$$\begin{aligned}E(\widehat{\beta}_0) &= E\left[\Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right)\right] \\ &= \Phi^{-1}E\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right) \\ &\simeq \Phi^{-1}\left[\frac{E(\widehat{n}^{10})}{E(\widehat{n}^{10}) + E(\widehat{n}^{00})}\right] \\ &= \Phi^{-1}\left(\frac{n^{10}}{n^{10} + n^{00}}\right)\end{aligned}$$

et de même

$$\begin{aligned}E(\widehat{\beta}_1) &= E\left[\Phi^{-1}\left(\frac{\widehat{n}^{11}}{\widehat{n}^{11} + \widehat{n}^{01}}\right) - \Phi^{-1}\left(\frac{\widehat{n}^{10}}{\widehat{n}^{10} + \widehat{n}^{00}}\right)\right] \\ &\simeq \Phi^{-1}\left(\frac{n^{11}}{n^{11} + n^{01}}\right) - \Phi^{-1}\left(\frac{n^{10}}{n^{10} + n^{00}}\right).\end{aligned}$$

En notant

$$\widehat{\beta} = \begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{pmatrix} \implies V(\widehat{\beta}) \equiv \Omega_{\widehat{\beta}} = V\begin{pmatrix} \widehat{\beta}_0 \\ \widehat{\beta}_1 \end{pmatrix}$$

on obtient

$$\begin{aligned}\Omega_{\hat{\beta}} &= \begin{pmatrix} V(\hat{\beta}_0) & -V(\hat{\beta}_0) \\ -V(\hat{\beta}_0) & V(\hat{\beta}_1) \end{pmatrix} \\ &= V(\hat{\beta}_0) \begin{pmatrix} 1 & -1 \\ -1 & v_0^1 \end{pmatrix}, \quad v_0^1 = \frac{V(\hat{\beta}_1)}{V(\hat{\beta}_0)}\end{aligned}$$

où (Annexe 1.A.7.3)

$$\begin{aligned}V(\hat{\beta}_0) &\simeq (\Phi'(\beta_0))^{-2} V(\hat{q}), \\ V(\hat{\beta}_1) &\simeq \frac{V(\hat{q}) - [\Phi'(\beta_0) - 2\Phi'(\beta_1)] (\Phi'(\beta_0))^{-1} V(\hat{q})}{(\Phi'(\beta_1))^2}, \\ Cov(\hat{\beta}_0, \hat{\beta}_1) &= -V(\hat{\beta}_0).\end{aligned}$$

Les scalaires $\Phi'(\beta_0)$ et $\Phi'(\beta_1)$ peuvent être estimés respectivement par $\Phi'(\hat{\beta}_0)$ et $\Phi'(\hat{\beta}_1)$.

Le paramètre $\hat{\beta}_1$ est statistiquement non significatif pour un risque de première espèce $0 < \alpha < 1$ que si

$$\left[\hat{\beta}_1 - z_{\alpha/2} \left(V(\hat{\beta}_1) \right)^{1/2}; \hat{\beta}_1 + z_{\alpha/2} \left(V(\hat{\beta}_1) \right)^{1/2} \right] \ni \beta_1$$

où $z_{\alpha/2}$ est le quartile d'ordre $\alpha/2$ de la distribution normale centrée réduite car, pour n^{yx} suffisamment grand (Théorème de la Limite Centrale),

$$\left\{ \hat{\beta}_1 \left[V(\hat{\beta}_1) \right]^{-1/2} \right\} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

1.5 Les résultats du modèle eu égard à la littérature

Les modèles utilisés dans le cadre des méthodes de capture recapture dans sa version multiliste, sont de plus en plus élaborés (Bohning et al., 2018; Borchers et al., 2002). Pour améliorer l'estimation de la taille des populations, sont exploités des modèles sophistiqués avec des covariables tels par exemple les modèles à variables latentes, des modèles

basés sur l'approche bayésienne, des modèles de comptage censurés etc. Cependant, la problématique reste fondamentalement l'estimation de la taille d'une population difficile à joindre.

Par ailleurs, aucun des travaux repertoriés jusqu'ici ne s'est préoccupé d'évaluer la pertinence éventuelle d'un estimateur naïf. De plus, dans le cadre des travaux réalisés à partir de deux occasions (sources) de capture, contrairement à notre démarche, la variance de l'estimateur de la taille de la population totale est estimée en supposant $n_{10}^{yx} + n_{11}^{yx}$ et $n_{01}^{yx} + n_{11}^{yx}$ constants. Non plus, il n'y a eu aucune recherche qui s'était donnée pour objectif principale de procéder à une analyse de régression. L'usage de plus en plus fréquent des méthodes de capture recapture et les données qui en découlent, suscitent de l'intérêt pour pousser un peu plus loin une analyse inférentielle sur les populations difficiles enquêtées.

1.6 Synthèse des principaux résultats et conclusion

Le développement de ce modèle dans le cadre de deux occasions de capture (sources) a permis de passer en revue les principaux estimateurs disponibles à ce jour pour estimer la taille inconnue d'une population difficile à joindre. L'estimateur de Lincoln-Petersen a été modifié et amélioré par des estimateurs tels l'estimateur de Chapman, l'estimateur de Zimmerman et l'estimateur de Chao. En faisant l'hypothèse de constance de $n_{10}^{yx} + n_{11}^{yx}$ et $n_{01}^{yx} + n_{11}^{yx}$, la variance de l'estimateur de la population totale \widehat{n}^{yx} a été estimée. Les chercheurs tels Chapman, Zimmerman et Chao ont dérivé d'autres expressions de la variance de \widehat{n}^{yx} en adoptant d'autres hypothèses.

Après avoir calculé la variance de l'estimateur de Lincoln-Petersen dans le cas où la population est stratifiée et comparé cette variance à celle du cas non stratifié, les probabilités conditionnelles \widehat{q} et $\widehat{\underline{q}}$ et leur correspondant naïf \widetilde{q} et $\widetilde{\underline{q}}$ ont été estimés. En relaxant l'hypothèse de constance de $n_{10}^{yx} + n_{11}^{yx}$ et $n_{01}^{yx} + n_{11}^{yx}$, la variance de l'estimateur de la population totale \widehat{n}^{yx} a été ré-calculée. La variance de la taille de la population totale observée o^{yx} a été aussi estimée sur la base d'une distribution trinomiale conditionnelle. Bien que les estimateurs naïfs sont en général asymptotiquement biaisés, leur variance est inférieure à la variance des estimateurs dérivés prenant en compte un estimateur \widehat{n}_{00}^{yx} de la

partie de la population qui n'a pas été observée. Un test d'absence de biais de l'estimateur naïf a été proposé. Sur la base de la variance de \widehat{n}^{yx} ré-calculée et la variance de o^{yx} , les estimateurs de la variance des probabilités conditionnelles \widehat{q} et \underline{q} et leur correspondant naïf \widetilde{q} et $\underline{\widetilde{q}}$ ont été estimés. De ces estimateurs des probabilités conditionnelles et leur variance, ont été estimés les paramètres du modèle binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$.

Les principales contributions de ce chapitre se résument en : a) Le calcul de la variance de l'estimateur de Lincoln-Petersen dans le cas où la population est stratifiée et la comparaison de cette variance à celle du cas non stratifié ; b) la définition et la caractérisation (biais, biais asymptotique, variance) d'un estimateur naïf pour l'estimateur des probabilités conditionnelles ; c) le calcul de la variance de \widehat{n}^{yx} et celle de \widehat{n}_{00}^{yx} en relaxant l'hypothèse de constance de $n_{10}^{yx} + n_{11}^{yx}$ et $n_{01}^{yx} + n_{11}^{yx}$; d) l'estimation des paramètres d'un petit modèle de régression binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$.

Les données issues de l'usage de plus en plus important des techniques de capture recapture dans leur version multiliste, suscitent l'élaboration et/ou l'adaption progressives de méthodes statistiques inférentielles devant exploiter ces données. L'idée c'est d'arriver dans un premier temps à généraliser les résultats de ce modèle à un nombre K quelconque de variables indépendantes binaires ou discrètes.

Bibliographie

- BOHNING, D. et AL. (2018). *Capture recapture methods for the social and medical sciences*. Boca Raton : CRC Press.
- BORCHERS, D. L. et AL. (2002). *Estimating animal abundance : closed populations*, volume 13. Springer Science and Business Media.
- BRITAIN, S. et BOHNING, D. (2009). Estimators in capture recapture studies with two sources. *AStA Advances in Statistical Analysis*, 93(1):23–47.
- BUCKLAND, S. T., ROWLEY, I. et & WILLIAMS, D. A. (1983). Estimation of survival from repeated sightings of tagged galahs. *The Journal of Animal Ecology*, pages 563–573.
- CHAO, A. (1987). Estimating the population size for capture recapture data with unequal catchability. *Biometrics*, pages 783–791.
- CHAO, A. (1989). Estimating population size for sparse data in capture recapture experiments. *Biometrics*, pages 427–438.
- CHAO, A. (2001). An overview of closed capture recapture models. *Journal of Agricultural, Biological, and Environmental Statistics*, 6(2):158–175.
- CHAO, A. (2014). Capture recapture for human populations. *Wiley StatsRef : Statistics Reference Online*, pages 1–16.
- CHAO, A. et AL. (2001). The applications of capture recapture models to epidemiological data. *Statistics in medicine*, 20:3123–3157.
- CHAO, A. et TSAY, P. K. (1998). A sample coverage approach to multiple system estimation with application to census undercount. *Journal of the American Statistical Association*, 93(441):283–293.
- CHAO, A. ; Pan, H. Y. et CHIANG, S. C. (2008). The petersen lincoln estimator and its extension to estimate the size of a shared population. *Biometrical Journal : Journal of Mathematical Methods in Biosciences*, 50(6):957–970.

- CHATTERJEE, K. (2016). *Some contributions to the analysis of dual-record system for estimating human population size*. Thèse de doctorat, Indian Statistical Institute-Kolkata.
- EL ALLAKI, F.; Christensen, J. et VALLIÈRES, A. (2015). Comparing capture recapture methods for estimation of the size of small and medium sized populations using empirical data on commercial turkey farms in canada. *Preventive veterinary medicine*, 120(1):86–95.
- GOURIEROUX, C. (1989). *Econométrie des variables qualitatives, 2ème édition*. Economica, Paris.
- HECKEL, D. G. et ROUGHGARDEN, J. (1979). A technique for estimating the size of lizard populations. *Ecology*, 60(5):966–975.
- IWGDMF (1995). Capture recapture and multiple record systems estimation i : history and theoretical development. international working group for disease monitoring and forecasting (iwgdmf). *American Journal of Epidemiology*, 142(10):1047–1058.
- KHAN, S. I.; Bhuiya, A. et UDDIN, A. J. (2004). Application of the capture recapture method for estimating number of mobile male sex workers in a port city of bangladesh. *Journal of health, population and nutrition*, pages 19–26.
- KROLL, N. E. (1989). Testing independence in 2 by 2 contingency tables. *Journal of Educational Statistics*, 14(1):47–79.
- LAVALLÉE, P. et RIVEST, L.-P. (2013). Capture recapture sampling and indirect sampling. *Quality control and applied statistics*, 58(5):467–468.
- LE CREN, E. D. (1965). A note on the history of mark recapture population estimates. *Journal of animal ecology*, 34:453–454.
- MANRIQUE-VALLIER, D. et AL. (2013). Multiple systems estimation techniques for estimating casualties in armed conflicts. *Counting Civilian Casualties*, 165.

- MITCHELL, S. A. (2014). *Capture recapture Estimation for Conflict Data and Hierarchical Models for Program Impact Evaluation*. Doctoral dissertation. Thèse de doctorat, Harvard University.
- MITTELHAMMER, R. C. (2013). *Mathematical statistics for economics and business*, volume 78. New York Springer.
- OTIS, D. L., BURNHAM, K. P., WHITE, G. C. et ANDERSON, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife monographs*, (62):3–135.
- RIVEST, L.-P. (2013). Théorie et applications des modèles capture recapture. *Bulletin AMQ*, vol. III, no 2, pp. 65-78.
- SCHNABEL, Z. E. (1938). The estimation of the total fish population of a lake. *The American Mathematical Monthly*, 45(6):348–352.
- SEBER, G. (1985). Approximate unbiased estimation in the multi sample single recapture census. *In Statistics in Ornithology*, pages 355–362. Springer.
- SEBER, G. et MANLY, B. (1985). Approximately unbiased variance estimation for the jolly seber mark recapture model : population size. *In Statistics in Ornithology*, pages 363–371. Springer.
- SEBER, G. A. (1986). A review of estimating animal abundance. *Biometrics*, pages 267–292.
- SEKAR, C. C. et DEMING, W. E. (1949). On a method of estimating birth and death rates and the extent of registration. *Journal of the American Statistical Association*, 44(245):101–115.
- SULEWSKI, P. (2019). Some contributions to practice of 2 by 2 contingency tables. *Journal of Applied Statistics*, 46(8):1438–1455.
- TOURANGEAU, R. E. et AL. (2014). *Hard-to-survey populations*. Cambridge University Press.

- TSAY, P. K. et CHAO, A. (2001). Population size estimation for capture recapture models with applications to epidemiological data. *Journal of Applied Statistics*, 28(1):25–36.
- VERGNE, T. (2012). *Les méthodes de capture recapture pour évaluer les systèmes de surveillance des maladies animales*. Thèse de doctorat, Université Paris Sud-Paris XI.
- WITTES, J. et SIDEL, V. W. (1968). A generalization of the simple capture-recapture model with applications to epidemiological research. *Journal of chronic diseases*, 21(5): 287–301.
- XU, C.; Sun, D. et HE, C. (2014). Objective bayesian analysis for a capture recapture model. *Annals of the Institute of Statistical Mathematics*, 66(2):245–278.
- YATES, F. (1984). Tests of significance for 2 by 2 contingency tables. *Journal of the Royal Statistical Society : Series A*, 147(3):426–449.
- ZASLAVSKY, A. M. et WOLFGANG, G. S. (1993). Triple system modeling of census, post enumeration survey, and administrative list data. *Journal of Business and Economic Statistics*, 11(3):279–288.
- ZWANE, E. et VAN DER HEIJDEN, P. (2008). Capture recapture studies with incomplete mixed categorical and continuous covariates. *Journal of data science*, 6:557–572.

Appendix

1.A Annexes

1.A.1 L'expression de l'estimateur \widehat{n}^{yx}

La taille de la sous population de caractéristique yx est donnée par :

$$\begin{aligned}\widehat{n}^{yx} &= n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx} + \widehat{n}_{00}^{yx} \\ &= n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx} + \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} \\ &= \frac{1}{n_{11}^{yx}} [n_{01}^{yx} n_{10}^{yx} + (n_{01}^{yx} + n_{10}^{yx} + n_{11}^{yx}) n_{11}^{yx}] \\ &= \frac{1}{n_{11}^{yx}} [n_{10}^{yx} (n_{01}^{yx} + n_{11}^{yx}) + (n_{10}^{yx} + n_{11}^{yx}) n_{11}^{yx}] \\ &= \frac{1}{n_{11}^{yx}} (n_{01}^{yx} + n_{11}^{yx}) (n_{10}^{yx} + n_{11}^{yx})\end{aligned}\tag{1.58}$$

1.A.2 Variance asymptotique de \widehat{n}^{yx} établie par Sekar et Deming

La formule de cette variance asymptotique a été obtenue par Sekar et Deming (1949) en supposant $n_{11}^{yx} + n_{10}^{yx}$, $n_{11}^{yx} + n_{01}^{yx}$ et n_{11}^{yx} constants. De cette hypothèse et en remarquant que $n_{11}^{yx} \sim B(n^{yx}, r_1^{yx} r_2^{yx})$ où $E(n_{11}^{yx}) = n^{yx} r_1^{yx} r_2^{yx}$ et $V(n_{11}^{yx}) \simeq n^{yx} r_1^{yx} r_2^{yx} (1 - r_1^{yx}) (1 - r_2^{yx})$ (probabilités à peu près identiques pour les deux sources), on obtient :

$$V(\widehat{n}^{yx}) = (n_{11}^{yx} + n_{10}^{yx})^2 (n_{11}^{yx} + n_{01}^{yx})^2 V\left(\frac{1}{n_{11}^{yx}}\right).$$

Par la méthode Delta (Mittelhammer, 2013), $V\left(\frac{1}{n^{yx}}\right) \approx \frac{V(n_{11}^{yx})}{[E(n_{11}^{yx})]^4}$ ce qui permet d'avoir

$$V(\widehat{n^{yx}}) \approx \widehat{n^{yx}} \frac{(1 - r_1^{yx})(1 - r_2^{yx})}{r_1^{yx} r_2^{yx}}.$$

Comme

$$\begin{aligned} \widehat{r_1^{yx}} &= \frac{n_{10}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}} = \frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \\ \widehat{r_2^{yx}} &= \frac{n_{01}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}} = \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \end{aligned}$$

Alors

$$\begin{aligned} V(\widehat{n^{yx}}) &\simeq \left(\frac{1}{n_{11}^{yx}}\right) (n_{11}^{yx} + n_{10}^{yx}) (n_{11}^{yx} + n_{01}^{yx}) \left[\frac{\left(1 - \frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) \left(1 - \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right)}{\left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right)} \right] \\ &\simeq \left(\frac{1}{n_{11}^{yx}}\right) (n_{11}^{yx} + n_{10}^{yx}) (n_{11}^{yx} + n_{01}^{yx}) \left[\frac{n_{10}^{yx} n_{01}^{yx}}{(n_{11}^{yx})^2} \right] \\ &\simeq \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{11}^{yx}}\right) \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{11}^{yx}}\right) \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \end{aligned}$$

ou encore

$$\begin{aligned} V(\widehat{n^{yx}}) &\simeq \left[\frac{(n_{11}^{yx})^2 + n_{01}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{10}^{yx} + n_{01}^{yx} n_{11}^{yx}}{(n_{11}^{yx})^2} \right] \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \\ &\simeq \left(1 + \frac{n_{01}^{yx} n_{10}^{yx}}{(n_{11}^{yx})^2} + \frac{n_{10}^{yx}}{n_{11}^{yx}} + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right) \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \\ &\simeq \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) + \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \frac{n_{01}^{yx} n_{10}^{yx}}{(n_{11}^{yx})^2} + \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \frac{n_{10}^{yx}}{n_{11}^{yx}} + \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) \frac{n_{01}^{yx}}{n_{11}^{yx}} \\ &\simeq \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right) + \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}}\right)^2 \left(\frac{1}{n_{11}^{yx}} + \frac{1}{n_{01}^{yx}} + \frac{1}{n_{10}^{yx}}\right) \\ &\simeq \widehat{n_{00}^{yx}} + \left(\frac{1}{n_{10}^{yx}} + \frac{1}{n_{01}^{yx}} + \frac{1}{n_{11}^{yx}}\right) (\widehat{n_{00}^{yx}})^2 \end{aligned} \tag{1.59}$$

1.A.3 Cas d'une population à deux états

On suppose que l'hétérogénéité de la population se matérialise par l'observation de deux états distinctifs \bar{z} et \underline{z} . On se propose alors d'estimer la taille de la population totale

n^{yx} telle que $n^{yx} = n^{yx}(\bar{z}) + n^{yx}(z)$ où $n^{yx}(\bar{z})$ et $n^{yx}(z)$ représentent la taille respective des sous-populations correspondant aux états \bar{z} et z . En référence aux équation (1.16), (1.17) et (1.18), le biais du à la non prise en compte de l'hétérogénéité est définie par :

$$b(n^{yx}) = \frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} - \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)}$$

ce qui se ré-écrit, en admettant que $n^{yx} = n^{yx}(\bar{z}) + n^{yx}(z)$, sous la forme

$$\begin{aligned} b(n^{yx}) &= \frac{(n_{10}^{yx}(\bar{z}) + n_{10}^{yx}(z))(n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(z))}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} - \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \\ &= n_{10}^{yx}(z) \left(\frac{n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) + n_{10}^{yx}(\bar{z}) \left(\frac{n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) \\ &\quad - \frac{n_{01}^{yx}(\bar{z}) n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(z) n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \\ &= n_{10}^{yx}(\bar{z}) \left(\frac{n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} - \frac{n_{01}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} \right) - n_{10}^{yx}(z) \left(\frac{n_{01}^{yx}(z)}{n_{11}^{yx}(z)} - \frac{n_{01}^{yx}(\bar{z}) + n_{01}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) \\ &= n_{10}^{yx}(\bar{z}) \left(\frac{n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})(n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z))} \right) \\ &\quad - n_{10}^{yx}(z) \left(\frac{n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(z)(n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z))} \right) \\ &= \left(\frac{n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) \left(\frac{n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \right) \\ &= \left(\frac{n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) \left(\frac{n_{10}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{10}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)} \right). \end{aligned}$$

Ce qui donne enfin de compte

$$\begin{aligned} b(n^{yx}) &= - \frac{[n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})][n_{10}^{yx}(z) n_{11}^{yx}(\bar{z}) - n_{10}^{yx}(\bar{z}) n_{11}^{yx}(z)]}{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z) (n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z))} \\ &= - \left(\frac{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \right) \left(\frac{n_{01}^{yx}(\bar{z}) n_{11}^{yx}(z) - n_{01}^{yx}(z) n_{11}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)} \right) \\ &\quad \times \left(\frac{n_{10}^{yx}(z) n_{11}^{yx}(\bar{z}) - n_{10}^{yx}(\bar{z}) n_{11}^{yx}(z)}{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)} \right) \\ &= - \frac{n_{11}^{yx}(\bar{z}) n_{11}^{yx}(z)}{n_{11}^{yx}(\bar{z}) + n_{11}^{yx}(z)} \left(\frac{n_{01}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{01}^{yx}(z)}{n_{11}^{yx}(z)} \right) \left(\frac{n_{10}^{yx}(\bar{z})}{n_{11}^{yx}(\bar{z})} - \frac{n_{10}^{yx}(z)}{n_{11}^{yx}(z)} \right) \end{aligned}$$

1.A.4 Variance asymptotique de \widehat{n}^{yx}

D'après le résultat de l'équation (1.58), on a :

$$\widehat{n}^{yx} = \frac{1}{n_{11}^{yx}} (n_{01}^{yx} + n_{11}^{yx}) (n_{10}^{yx} + n_{11}^{yx}).$$

Soit $\widehat{n}^{yx} \equiv f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx}) = \frac{1}{n_{11}^{yx}} (n_{01}^{yx} + n_{11}^{yx}) (n_{10}^{yx} + n_{11}^{yx})$. Une approximation de Taylor à l'ordre un au voisinage de

$$\mu = (E(n_{01}^{yx}), E(n_{10}^{yx}), E(n_{11}^{yx}))$$

de l'expression de \widehat{n}^{yx} est donnée par :

$$\begin{aligned} \widehat{n}^{yx} &\simeq f[E(n_{01}^{yx}), E(n_{10}^{yx}), E(n_{11}^{yx})] \\ &+ [n_{01}^{yx} - E(n_{01}^{yx})] \left(\frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{01}^{yx}} \Big|_{\mu} \right) \\ &+ [n_{10}^{yx} - E(n_{10}^{yx})] \left(\frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{10}^{yx}} \Big|_{\mu} \right) \\ &+ [n_{11}^{yx} - E(n_{11}^{yx})] \left(\frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{11}^{yx}} \Big|_{\mu} \right) \end{aligned}$$

où

$$\begin{aligned}
\left. \frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{01}^{yx}} \right|_{\mu} &\equiv c_{01} = \frac{E(n_{10}^{yx}) + E(n_{11}^{yx})}{E(n_{11}^{yx})} = \frac{1}{r_2^{yx}} \\
\left. \frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{10}^{yx}} \right|_{\mu} &\equiv c_{10} = \frac{E(n_{01}^{yx}) + E(n_{11}^{yx})}{E(n_{11}^{yx})} = \frac{1}{r_1^{yx}} \\
\left. \frac{\partial f(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{11}^{yx}} \right|_{\mu} &\equiv c_{11} = \frac{[(E(n_{10}^{yx}) + E(n_{11}^{yx})) + (E(n_{01}^{yx}) + E(n_{11}^{yx}))][E(n_{11}^{yx})]}{[E(n_{11}^{yx})]^2} \\
&\quad - \frac{[(E(n_{01}^{yx}) + E(n_{11}^{yx}))(E(n_{10}^{yx}) + E(n_{11}^{yx}))]}{[E(n_{11}^{yx})]^2} \\
&= \frac{[E(n_{11}^{yx})]^2 - E(n_{10}^{yx})E(n_{01}^{yx})}{[E(n_{11}^{yx})]^2} \\
&= \frac{(r_1^{yx} r_2^{yx})^2 - [r_1^{yx}(1 - r_2^{yx})(1 - r_1^{yx})r_2^{yx}]}{(r_1^{yx} r_2^{yx})^2} \\
&= \frac{r_1^{yx} + r_2^{yx} - 1}{(r_1^{yx} r_2^{yx})}
\end{aligned}$$

L'espérance mathématique est approximativement égale à

$$\begin{aligned}
E(\widehat{n^{yx}}) &\simeq \frac{1}{E(n_{11}^{yx})} [E(n_{01}^{yx}) + E(n_{11}^{yx})][E(n_{10}^{yx}) + E(n_{11}^{yx})] \\
&\simeq \widehat{n^{yx}}.
\end{aligned}$$

Sous l'hypothèse que $cov(n_{\omega_1}^{yx}, n_{\omega_2}^{yx}) = 0$, par la *méthode delta*, variance approximée est telle que

$$\begin{aligned}
V(\widehat{n^{yx}}) &\simeq c_{11}^2 V(n_{11}^{yx}) + c_{10}^2 V(n_{10}^{yx}) + c_{01}^2 V(n_{01}^{yx}) \\
&\simeq \widehat{n^{yx}} \left\{ \left(\frac{1 - r_1^{yx} - r_2^{yx}}{r_1^{yx} r_2^{yx}} \right)^2 r_1^{yx} r_2^{yx} (1 - r_1^{yx} r_2^{yx}) \right. \\
&\quad + \left(\frac{1}{r_1^{yx}} \right)^2 r_1^{yx} (1 - r_2^{yx}) [1 - r_1^{yx} (1 - r_2^{yx})] \\
&\quad \left. + \left(\frac{1}{r_2^{yx}} \right)^2 (1 - r_1^{yx}) r_2^{yx} [1 - (1 - r_1^{yx}) r_2^{yx}] \right\}
\end{aligned}$$

L'expression approximée de la variance se ré-écrit :

$$V(\widehat{n^{yx}}) \simeq \frac{\widehat{n^{yx}}}{\widehat{r_1^{yx}} \widehat{r_2^{yx}}} \left\{ (1 - r_1^{yx} - r_2^{yx})^2 (1 - r_1^{yx} r_2^{yx}) \right. \\ \left. + r_2^{yx} (1 - r_2^{yx}) [1 - r_1^{yx} (1 - r_2^{yx})] \right. \\ \left. + r_1^{yx} (1 - r_1^{yx}) [1 - (1 - r_1^{yx}) r_2^{yx}] \right\}$$

Par suite

$$\frac{\widehat{r_1^{yx}} \widehat{r_2^{yx}}}{\widehat{n^{yx}}} V(\widehat{n^{yx}}) \simeq (1 - (r_1^{yx} + r_2^{yx}))^2 (1 - r_1^{yx} r_2^{yx}) \\ + r_1^{yx} (1 - r_1^{yx}) + r_2^{yx} (1 - r_2^{yx}) \\ - r_1^{yx} r_2^{yx} \left[(1 - r_1^{yx})^2 + (1 - r_2^{yx})^2 \right] \\ = r_1^{yx} (1 - r_1^{yx}) + r_2^{yx} (1 - r_2^{yx}) + (1 - (r_1^{yx} + r_2^{yx}))^2 \\ - r_1^{yx} r_2^{yx} \left[(1 - r_1^{yx})^2 + (1 - r_2^{yx})^2 + (1 - (r_1^{yx} + r_2^{yx}))^2 \right].$$

En se rappelant que

$$\widehat{r_1^{yx}} = \frac{n_{10}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}} = \frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}, \\ \widehat{r_2^{yx}} = \frac{n_{01}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}} = \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}, \\ 1 - (\widehat{r_1^{yx}} + \widehat{r_2^{yx}}) = \frac{(n_{10}^{yx} + n_{11}^{yx})(n_{01}^{yx} + n_{11}^{yx}) - n_{11}^{yx}(n_{01}^{yx} + n_{11}^{yx} + n_{10}^{yx} + n_{11}^{yx})}{(n_{10}^{yx} + n_{11}^{yx})(n_{01}^{yx} + n_{11}^{yx})} \\ = \frac{n_{01}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + (n_{11}^{yx})^2 - (n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + 2(n_{11}^{yx})^2)}{(n_{10}^{yx} + n_{11}^{yx})(n_{01}^{yx} + n_{11}^{yx})} \\ = \frac{n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2}{(n_{10}^{yx} + n_{11}^{yx})(n_{01}^{yx} + n_{11}^{yx})} \\ \frac{\widehat{r_1^{yx}} \widehat{r_2^{yx}}}{\widehat{n^{yx}}} = \frac{\left(\frac{n_{10}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}}\right) \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{\widehat{n^{yx}}}\right)}{\widehat{n^{yx}}} = \frac{(n_{10}^{yx} + n_{11}^{yx})(n_{01}^{yx} + n_{11}^{yx})}{\widehat{n^{yx}}} \frac{1}{(\widehat{n^{yx}})^2} \\ = \frac{(n_{11}^{yx})^3}{(n_{01}^{yx} + n_{11}^{yx})^2 (n_{10}^{yx} + n_{11}^{yx})^2}$$

on obtient :

$$\begin{aligned}
\frac{\widehat{r}_1^{yx} \widehat{r}_2^{yx}}{\widehat{n}^{yx}} V(\widehat{n}^{yx}) &= \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(1 - \frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) + \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \left(1 - \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \\
&+ \left[1 - \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} + \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]^2 \\
&- \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \times \left\{ \left(1 - \frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right)^2 + \left(1 - \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right)^2 \right. \\
&+ \left. \left[1 - \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} + \frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]^2 \right\} \\
&= \frac{n_{01}^{yx} n_{11}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})^2} + \frac{n_{10}^{yx} n_{11}^{yx}}{(n_{10}^{yx} + n_{11}^{yx})^2} + \left[\frac{n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right]^2 \\
&- \frac{(n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \times \left\{ \left(\frac{n_{01}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right)^2 + \left(\frac{n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right)^2 \right. \\
&+ \left. \left[\frac{n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right]^2 \right\}
\end{aligned}$$

Donc

$$\begin{aligned}
\frac{\widehat{r}_1^{yx} \widehat{r}_2^{yx}}{\widehat{n}^{yx}} V(\widehat{n}^{yx}) &= \frac{n_{01}^{yx} n_{11}^{yx} (n_{10}^{yx} + n_{11}^{yx})^2 + (n_{01}^{yx} + n_{11}^{yx})^2 n_{10}^{yx} n_{11}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})^2 (n_{10}^{yx} + n_{11}^{yx})^2} \\
&+ \left[\frac{n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right]^2 \\
&- \frac{(n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})^2 (n_{10}^{yx} + n_{11}^{yx})^2} \times \left\{ \frac{[n_{01}^{yx} (n_{10}^{yx} + n_{11}^{yx})]^2 + [(n_{01}^{yx} + n_{11}^{yx}) n_{10}^{yx}]^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right. \\
&+ \left. \frac{[n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2]^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right\} \\
&= \frac{(n_{11}^{yx})^3}{(n_{01}^{yx} + n_{11}^{yx})^2 (n_{10}^{yx} + n_{11}^{yx})^2} \left\{ \frac{n_{01}^{yx} (n_{10}^{yx} + n_{11}^{yx})^2}{(n_{11}^{yx})^2} + \frac{(n_{01}^{yx} + n_{11}^{yx})^2 n_{10}^{yx}}{(n_{11}^{yx})^2} \right. \\
&+ \frac{[n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2]^2}{(n_{11}^{yx})^3} - \frac{1}{n_{11}^{yx}} \left[\frac{[n_{01}^{yx} (n_{10}^{yx} + n_{11}^{yx})]^2 + [(n_{01}^{yx} + n_{11}^{yx}) n_{10}^{yx}]^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right. \\
&+ \left. \left. \frac{[n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2]^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right] \right\}
\end{aligned}$$

Ainsi,

$$\begin{aligned} \frac{\widehat{r}_1^{yx} \widehat{r}_2^{yx}}{\widehat{n}^{yx}} V(\widehat{n}^{yx}) &= \frac{\widehat{r}_1^{yx} \widehat{r}_2^{yx}}{\widehat{n}^{yx}} \left\{ n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}} \right)^2 \right. \\ &\quad + \frac{1}{n_{11}^{yx}} \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} - 1 \right)^2 - \frac{1}{n_{11}^{yx}} \left[\frac{n_{01}^{yx} (n_{10}^{yx} + n_{11}^{yx})}{n_{01}^{yx} + n_{11}^{yx}} + \frac{(n_{01}^{yx} + n_{11}^{yx}) n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right] \\ &\quad \left. - \frac{1}{n_{11}^{yx}} \frac{[n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2]^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right\} \\ V(\widehat{n}^{yx}) &= n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}} \right)^2 \\ &\quad + \frac{1}{n_{11}^{yx}} \left(\frac{n_{01}^{yx} n_{10}^{yx}}{n_{11}^{yx}} - 1 \right)^2 - \frac{1}{n_{11}^{yx}} \left[\frac{n_{01}^{yx} (n_{10}^{yx} + n_{11}^{yx})}{n_{01}^{yx} + n_{11}^{yx}} + \frac{(n_{01}^{yx} + n_{11}^{yx}) n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right] \\ &\quad - \frac{[n_{01}^{yx} n_{10}^{yx} - (n_{11}^{yx})^2]^2}{\frac{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})}{n_{11}^{yx}}} \end{aligned}$$

et donc

$$\begin{aligned} V(\widehat{n}^{yx}) &\simeq n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}} \right)^2 + \frac{1}{n_{11}^{yx}} \left(\widehat{n}_{00}^{yx} - 1 \right)^2 \\ &\quad - \left\{ \frac{(\widehat{n}_{00}^{yx} - 1)^2}{\widehat{n}^{yx}} + \frac{1}{n_{11}^{yx}} \left[n_{01}^{yx} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) + n_{10}^{yx} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right] \right\} \quad (1.60) \end{aligned}$$

1.A.5 Biais asymptotique de l'estimateur naïf de \widetilde{q}

Démonstration. Dans un premier temps, on exprime o^{yx} , n^{yx} , \widehat{n}^{yx} , n_{ω}^{yx} et $\widehat{n}_{\omega}^{yx}$ en fonction des fréquences définies à partir du modèle statistique sous-jacent décrit en (1.10) et en

utilisant aussi le tableau (1.2) :

$$\begin{aligned}
n^{11} &= \bar{q}\dot{q}N; & n^{01} &= (1 - \bar{q})\dot{q}N; \\
n^{10} &= \underline{q}(1 - \dot{q})N; & n^{00} &= (1 - \dot{q})(1 - \underline{q})N; \\
n_{00}^{yx} &= n^{yx}(1 - r_1^{yx})(1 - r_2^{yx}); & n_{11}^{yx} &= n^{yx}(r_1^{yx})(r_2^{yx}); \\
n_{01}^{yx} &= n^{yx}(1 - r_1^{yx})(r_2^{yx}); & n_{10}^{yx} &= n^{yx}(r_1^{yx})(1 - r_2^{yx}); \\
n_{00}^{11} &= (1 - r_1^{11})(1 - r_2^{11})\bar{q}\dot{q}N; & n_{01}^{11} &= (1 - r_1^{11})(r_2^{11})\bar{q}\dot{q}N; \\
n_{10}^{11} &= (r_1^{11})(1 - r_2^{11})\bar{q}\dot{q}N; & n_{11}^{11} &= (r_1^{11})(r_2^{11})\bar{q}\dot{q}N; \\
n_{00}^{01} &= (1 - r_1^{01})(1 - r_2^{01})N\underline{q}(1 - \dot{q}); & n_{01}^{01} &= (1 - r_1^{01})(r_2^{01})N\underline{q}(1 - \dot{q}); \\
n_{10}^{01} &= (r_1^{10})(1 - r_2^{10})N\underline{q}(1 - \dot{q}); & n_{11}^{01} &= (r_1^{10})(r_2^{10})N\underline{q}(1 - \dot{q}),
\end{aligned}$$

Ensuite, en remplaçant les fréquences précédemment déterminées dans l'expression de $b_{\bar{q}}$ établie en (1.37) et en passant à la limite quand $N \rightarrow \infty$, on obtient :

$$\begin{aligned}
\lim_{N \rightarrow \infty} b_{\bar{q}} &= \lim_{N \rightarrow \infty} \left[\frac{N^2}{[\bar{q}\dot{q} + \underline{q}(1 - \dot{q})] N^2} \right] \\
&\times \lim_{N \rightarrow \infty} \left[\frac{\bar{q}\dot{q}\underline{q}(1 - \dot{q})}{\bar{q}\dot{q} + \underline{q}(1 - \dot{q}) + (1 - r_2^{01})(1 - r_1^{01})\underline{q}(1 - \dot{q}) + (1 - r_2^{11})(1 - r_1^{11})\bar{q}\dot{q}} \right] \\
&\times \lim_{N \rightarrow \infty} [(1 - r_2^{01})(1 - r_1^{01}) - (1 - r_1^{11})(1 - r_2^{11})]
\end{aligned}$$

Ce qui donne finalement

$$\lim_{N \rightarrow \infty} b_{\bar{q}} = \frac{\bar{q}\dot{q}\underline{q}(1 - \dot{q}) [(1 - r_2^{01})(1 - r_1^{01}) - (1 - r_1^{11})(1 - r_2^{11})]}{[\bar{q}\dot{q} + \underline{q}(1 - \dot{q})] \{ \bar{q}\dot{q} [1 + (1 - r_2^{11})(1 - r_1^{11})] + \underline{q}(1 - \dot{q}) [1 + (1 - r_2^{01})(1 - r_1^{01})] \}}.$$

Par suite, $b_{\bar{q}}$ est asymptotiquement nul si et seulement si

$$(1 - r_1^{11})(1 - r_2^{11}) = (1 - r_1^{01})(1 - r_2^{01}). \quad (1.61)$$

En d'autres termes, on retrouve le fait que, pour que l'estimateur naïf de \bar{q} soit (asymptotiquement) sans biais, il faut que la probabilité que les individus ne fassent pas partie d'aucune des deux listes soit identique pour les individus de caractéristiques $yx = 01$ et

ceux de caractéristiques $yx = 11$ - c'est à dire pour tous les individus de caractéristique $x = 1$. □

1.A.6 Espérance et Variance de o^{yx}

1.A.6.1 Espérance mathématique de o^{yx}

Distribution du triplet $(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})$ étant donné n_{00}^{yx} . On sait que

$$(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}, n_{00}^{yx}) \sim \mu(n^{yx}, p_{11}^{yx}, p_{10}^{yx}, p_{01}^{yx}, p_{00}^{yx})$$

et donc

$$\begin{aligned} p(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}, n_{00}^{yx}) &= \frac{n^{yx}!}{n_{11}^{yx}! n_{10}^{yx}! n_{01}^{yx}! n_{00}^{yx}!} [r_1^{yx} r_2^{yx}]^{n_{11}^{yx}} [r_1^{yx} (1 - r_2^{yx})]^{n_{10}^{yx}} \\ &\quad \times [(1 - r_1^{yx}) r_2^{yx}]^{n_{01}^{yx}} [(1 - r_1^{yx}) (1 - r_2^{yx})]^{n_{00}^{yx}}. \end{aligned}$$

Les distributions marginales sont telles que

$$n_{\omega}^{yx} \sim \mathcal{B}(n^{yx}, p_{\omega}^{yx}), \quad \omega \in \{11, 10, 01, 00\}.$$

La distribution conditionnelle de $(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx})$ sachant n_{00}^{yx} est telle que

$$p(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx} / n_{00}^{yx}) = \frac{p(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}, n_{00}^{yx})}{p(n_{00}^{yx})}.$$

On sait que $n_{00}^{yx} \sim \mathcal{B}(n^{yx}, p_{00}^{yx})$ où $p_{00}^{yx} = (1 - r_1^{yx})(1 - r_2^{yx})$ et donc

$$\begin{aligned}
p(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}/n_{00}^{yx}) &= \frac{\frac{n^{yx}!}{n_{11}^{yx}!n_{10}^{yx}!n_{01}^{yx}!n_{00}^{yx}!}}{n^{yx}!} [r_1^{yx} r_2^{yx}]^{n_{11}^{yx}} [r_1^{yx} (1 - r_2^{yx})]^{n_{10}^{yx}} \\
&\times \frac{[(1 - r_1^{yx}) r_2^{yx}]^{n_{01}^{yx}} [(1 - r_1^{yx})(1 - r_2^{yx})]^{n_{00}^{yx}}}{[(1 - r_1^{yx})(1 - r_2^{yx})]^{n_{00}^{yx}} [1 - (1 - r_1^{yx})(1 - r_2^{yx})]^{n^{yx} - n_{00}^{yx}}} \\
&= \frac{(n^{yx} - n_{00}^{yx})!}{n_{11}^{yx}!n_{10}^{yx}!n_{01}^{yx}!} [r_1^{yx} r_2^{yx}]^{n_{11}^{yx}} [r_1^{yx} (1 - r_2^{yx})]^{n_{10}^{yx}} \\
&\times \frac{[(1 - r_1^{yx}) r_2^{yx}]^{n_{01}^{yx}}}{[1 - (1 - r_1^{yx})(1 - r_2^{yx})]^{n^{yx} - n_{00}^{yx}}} \\
&= \frac{(n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx})!}{n_{11}^{yx}!n_{10}^{yx}!n_{01}^{yx}!} [r_1^{yx} r_2^{yx}]^{n_{11}^{yx}} [r_1^{yx} (1 - r_2^{yx})]^{n_{10}^{yx}} \\
&\times [(1 - r_1^{yx}) r_2^{yx}]^{n_{01}^{yx}} [1 - (1 - r_1^{yx})(1 - r_2^{yx})]^{-n_{11}^{yx} - n_{10}^{yx} - n_{01}^{yx}}.
\end{aligned}$$

Par suite,

$$\begin{aligned}
p(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}/n_{00}^{yx}) &= \frac{(n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx})!}{n_{11}^{yx}!n_{10}^{yx}!n_{01}^{yx}!} \left[\frac{r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right]^{n_{11}^{yx}} \\
&\times \left[\frac{(1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right]^{n_{10}^{yx}} \left[\frac{(1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right]^{n_{01}^{yx}}.
\end{aligned}$$

Par conséquent,

$$(n_{11}^{yx}, n_{10}^{yx}, n_{01}^{yx}/n_{00}^{yx}) \sim \mu(\tilde{n}^{yx}, \tilde{p}_{11}^{yx}, \tilde{p}_{10}^{yx}, \tilde{p}_{01}^{yx})$$

où

$$\begin{aligned}
\tilde{n}^{yx} &= n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} \\
\tilde{p}_{\omega}^{yx} &= \frac{p_{\omega}^{yx}}{1 - p_{00}^{yx}}, \quad \omega \in \{11, 10, 01\}.
\end{aligned}$$

Ainsi, $(n_{\omega}^{yx}/n_{00}^{yx}) \sim B(\tilde{n}^{yx}, \tilde{p}_{\omega}^{yx})$ et donc $V(n_{\omega}^{yx}/n_{00}^{yx}) = \tilde{n}^{yx} \tilde{p}_{\omega}^{yx} (1 - \tilde{p}_{\omega}^{yx})$.

L'espérance mathématique de o^{yx} est telle que :

$$\begin{aligned}
E(o^{yx}) &\equiv E(\tilde{n}^{yx}) = E(n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx}) \\
&= E(n_{11}^{yx}) + E(n_{10}^{yx}) + E(n_{01}^{yx}) \\
&= \tilde{n}^{yx} \left[\frac{r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} + \frac{(1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right. \\
&\quad \left. + \frac{(1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \\
&= \tilde{n}^{yx} \left[\frac{r_1^{yx} r_2^{yx} + (1 - r_2^{yx}) r_1^{yx} + (1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right].
\end{aligned}$$

Soit

$$\begin{aligned}
E(\tilde{n}^{yx}) &= \tilde{n}^{yx} \left[\frac{r_1^{yx} + r_2^{yx} - r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \\
&= \tilde{n}^{yx}
\end{aligned} \tag{1.62}$$

1.A.6.2 Variance de o^{yx}

En supposant les $(n_{\omega}^{yx}/n_{00}^{yx})$ indépendantes, alors

$$\begin{aligned}
V(o^{yx}) &\equiv V(\tilde{n}^{yx}) = V(n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx}) \\
&= V(n_{11}^{yx}) + V(n_{10}^{yx}) + V(n_{01}^{yx}) \\
&= \tilde{n}^{yx} \left\{ \left[\frac{r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[1 - \frac{r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \right. \\
&\quad + \left[\frac{(1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[1 - \frac{(1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \\
&\quad \left. + \left[\frac{(1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[1 - \frac{(1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \right\}
\end{aligned}$$

ou encore

$$\begin{aligned}
V(o^{yx}) \equiv V(\tilde{n}^{yx}) &= \tilde{n}^{yx} \left\{ \left[\frac{r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[\frac{[1 - (1 - r_1^{yx})(1 - r_2^{yx})] - r_1^{yx} r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \right. \\
&+ \left[\frac{(1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[\frac{[1 - (1 - r_1^{yx})(1 - r_2^{yx})] - (1 - r_2^{yx}) r_1^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \\
&+ \left. \left[\frac{(1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \left[\frac{[1 - (1 - r_1^{yx})(1 - r_2^{yx})] - (1 - r_1^{yx}) r_2^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] \right\} \\
&= \left[\frac{\tilde{n}^{yx}}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \right] [r_1^{yx} r_2^{yx} (r_1^{yx} + r_2^{yx} - 2r_1^{yx} r_2^{yx}) \\
&+ r_1^{yx} (1 - r_2^{yx}) r_2^{yx} + r_1^{yx} (1 - r_1^{yx}) r_2^{yx}] \\
&= \frac{2r_1^{yx} r_2^{yx} (1 - r_1^{yx} r_2^{yx})}{1 - (1 - r_1^{yx})(1 - r_2^{yx})} \tilde{n}^{yx}.
\end{aligned}$$

Par suite,

$$\begin{aligned}
V(o^{yx}) \equiv V(\tilde{n}^{yx}) &= \frac{2 \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \left[1 - \left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]}{1 - \left[\left(\frac{n_{01}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]} \tilde{n}^{yx} \\
&= \frac{2 \left[\frac{(n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right] \left[\frac{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx}) - (n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right]}{\frac{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx}) - n_{10}^{yx} n_{01}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})}} \tilde{n}^{yx} \\
&= \frac{2 \left[\frac{(n_{11}^{yx})^2}{(n_{11}^{yx})^2 + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}} \right] [n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}]}{(n_{11}^{yx})^2 + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx}} \tilde{n}^{yx} \\
&= \frac{2 \left[\frac{(n_{11}^{yx})^2}{(n_{11}^{yx})^2 + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}} \right] [n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}]}{n_{11}^{yx}} \\
&= 2n_{11}^{yx} \left\{ \frac{(n_{11}^{yx})^2 + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}}{(n_{11}^{yx})^2 + n_{11}^{yx} n_{10}^{yx} + n_{11}^{yx} n_{01}^{yx} + n_{10}^{yx} n_{10}^{yx}} \right\} \\
&- 2n_{11}^{yx} \left[\frac{(n_{11}^{yx})^2}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right].
\end{aligned}$$

Ainsi,

$$V(o^{yx}) \equiv V(\tilde{n}^{yx}) = 2n_{11}^{yx} \left\{ 1 - \left[\left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}} \right) \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right) \right]^{-1} \right\}. \quad (1.63)$$

1.A.6.3 Comparaison de $V(\widehat{\bar{q}})$ et de $V(\widetilde{\bar{q}})$.

Les variances de $V(\widehat{\bar{q}})$ et de $V(\widetilde{\bar{q}})$ s'écrivent respectivement (équations 1.34 et 1.55) :

$$V(\widehat{\bar{q}}) \simeq \frac{[E(\widehat{n^{01}})]^2}{[E(\widehat{n^{01}}) + E(\widehat{n^{11}})]^4} V(\widehat{n^{11}}) + \frac{[E(\widehat{n^{11}})]^2}{[E(\widehat{n^{01}}) + E(\widehat{n^{11}})]^4} V(\widehat{n^{01}}) \quad (1.64)$$

$$V(\widetilde{\bar{q}}) \simeq \frac{(o^{01})^2}{(o^{11} + o^{01})^4} V(o^{11}) + \frac{(o^{11})^2}{(o^{11} + o^{01})^4} V(o^{01}). \quad (1.65)$$

On rappelle que $n^{yx} = n_{11}^{yx} + n_{10}^{yx} + n_{01}^{yx} + n_{00}^{yx}$ ou encore $n^{yx} = o^{yx} + n_{00}^{yx}$. Pour alléger la notation, on note $a = o^{01}$, $b = o^{11}$, $c = E(\widehat{n_{00}^{01}})$ et $d = E(\widehat{n^{11}})$. Donc, $E(\widehat{n^{11}}) = o^{11} + E(\widehat{n^{11}}) = b + d$ et $E(\widehat{n^{01}}) = o^{01} + E(\widehat{n_{00}^{01}}) = a + c$. On procède en deux étapes.

Etape 1. On compare deux à deux les coefficients de l'équation (1.64) à ceux de l'équation (1.65). Par suite, les coefficients de l'équation (1.65) sont inférieurs à ceux de l'équation (1.64) si et seulement si

$$\begin{cases} \frac{a}{(a+b)^2} < \frac{a+c}{(a+c+b+d)^2} \\ \frac{b}{(a+b)^2} < \frac{b+d}{(a+c+b+d)^2} \end{cases} \iff \begin{cases} a(a+c+b+d)^2 < (a+c)(a+b)^2 \\ b(a+c+b+d)^2 < (b+d)(a+b)^2 \end{cases} \\ \iff \begin{cases} a[(a+c+b+d)^2 - (a+b)^2] < c(a+b)^2 \\ b[(a+c+b+d)^2 - (a+b)^2] < d(a+b)^2 \end{cases}.$$

En divisant membre à membre, on obtient $\frac{a}{b} < \frac{c}{d}$ c'est-à-dire les coefficients de $V(\widetilde{\bar{q}})$ sont inférieurs à ceux de $V(\widehat{\bar{q}})$ si et seulement si

$$\frac{o^{01}}{o^{11}} < \frac{E(\widehat{n_{00}^{01}})}{E(\widehat{n_{00}^{11}})}.$$

Etape 2. On compare les variances $V(\widehat{n^{yx}})$ et $V(o^{yx})$.

Les variances $V(\widehat{n^{yx}})$ et $V(o^{yx})$ obtenues respectivement aux équations (1.60) et (1.63)

sont données par :

$$V(\widehat{n}^{yx}) \simeq n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right)^2 + \frac{1}{n_{11}^{yx}} \left(\widehat{n}_{00}^{yx} - 1\right)^2 - \left\{ \frac{\left(\widehat{n}_{00}^{yx} - 1\right)^2}{\widehat{n}^{yx}} + \frac{1}{n_{11}^{yx}} \left[n_{01}^{yx} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) + n_{10}^{yx} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right] \right\},$$

$$V(o^{yx}) = 2n_{11}^{yx} \left\{ 1 - \left[\left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right) \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right) \right]^{-1} \right\}.$$

Evaluons $\Delta_V = V(\widehat{n}^{yx}) - V(o^{yx})$. Alors

$$\begin{aligned} \Delta_V &= n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right)^2 + n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right)^2 + \frac{1}{n_{11}^{yx}} \left(\widehat{n}_{00}^{yx} - 1\right)^2 \\ &\quad - \left\{ \frac{\left(\widehat{n}_{00}^{yx} - 1\right)^2}{\widehat{n}^{yx}} + \frac{1}{n_{11}^{yx}} \left[n_{01}^{yx} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) + n_{10}^{yx} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right] \right\} \\ &\quad - 2n_{11}^{yx} \left\{ 1 - \left[\left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right) \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right) \right]^{-1} \right\} \\ &= \underbrace{n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right)^2 - \frac{1}{n_{11}^{yx}} \left[n_{01}^{yx} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \right]}_{=A} \\ &\quad + \underbrace{n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right)^2 - \frac{1}{n_{11}^{yx}} \left[n_{10}^{yx} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]}_{=B} + \\ &\quad + \underbrace{\left[\left(\widehat{n}_{00}^{yx} - 1\right)^2 \left(\frac{1}{n_{11}^{yx}} - \frac{1}{\widehat{n}^{yx}} \right) \right]}_{=C} + \underbrace{2n_{11}^{yx} \left\{ \left[\left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right) \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right) \right]^{-1} - 1 \right\}}_{=D}. \end{aligned}$$

A l'évidence, les termes $A > 0$, $B > 0$ et $C > 0$ car respectivement

$$\begin{aligned} n_{01}^{yx} \left(1 + \frac{n_{10}^{yx}}{n_{11}^{yx}}\right)^2 &> \frac{n_{01}^{yx}}{n_{11}^{yx}} \left(\frac{n_{10}^{yx} + n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right), \\ n_{10}^{yx} \left(1 + \frac{n_{01}^{yx}}{n_{11}^{yx}}\right)^2 &> \frac{n_{10}^{yx}}{n_{11}^{yx}} \left(\frac{n_{01}^{yx} + n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right), \\ \frac{1}{n_{11}^{yx}} - \frac{1}{\widehat{n}^{yx}} &= \frac{\widehat{n}^{yx} - n_{11}^{yx}}{n_{11}^{yx} \widehat{n}^{yx}} > 0 \implies \frac{1}{n_{11}^{yx}} > \frac{1}{\widehat{n}^{yx}}. \end{aligned}$$

De même, le terme $D > 0$ car

$$\begin{aligned} & \left[\left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right] < 1 \\ \Rightarrow & \left[\left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}} \right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}} \right) \right]^{-1} > 1. \end{aligned}$$

En conséquence, $\Delta_V > 0$. Il s'ensuit que $V(\widehat{n^{yx}}) > V(o^{yx})$.

1.A.7 Espérance et Variance de $\widehat{n_{00}^{yx}}$

1.A.7.1 Espérance mathématique et variance de $\widehat{n_{00}^{yx}}$

L'espérance mathématique L'approximation par la méthode delta (Mittelhammer, 2013) permet d'écrire :

$$\begin{aligned} E(\widehat{n_{00}^{yx}}) & \simeq \frac{E(n_{01}^{yx}) E(n_{10}^{yx})}{E(n_{11}^{yx})} \\ & \simeq \frac{n^{yx} (1 - r_1^{yx}) r_2^{yx} n^{yx} r_1^{yx} (1 - r_2^{yx})}{n^{yx} r_1^{yx} r_2^{yx}} \\ & \simeq E(n_{00}^{yx}). \end{aligned}$$

et sous l'hypothèse d'indépendance, la variance par :

$$V(\widehat{n_{00}^{yx}}) \simeq (c_{01})^2 V(n_{01}^{yx}) + (c_{10})^2 V(n_{10}^{yx}) + (c_{11})^2 V(n_{11}^{yx})$$

où

$$\begin{aligned} c_{01} & = \left. \frac{\partial f_V(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{01}^{yx}} \right|_{\mu} = \left. \frac{n_{10}^{yx}}{n_{11}^{yx}} \right|_{\mu} = \frac{E(n_{10}^{yx})}{E(n_{11}^{yx})} = \frac{1 - r_2^{yx}}{r_2^{yx}} \\ c_{10} & = \left. \frac{\partial f_V(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{10}^{yx}} \right|_{\mu} = \left. \frac{n_{01}^{yx}}{n_{11}^{yx}} \right|_{\mu} = \frac{E(n_{01}^{yx})}{E(n_{11}^{yx})} = \frac{1 - r_1^{yx}}{r_1^{yx}} \\ c_{11} & = \left. \frac{\partial f_V(n_{01}^{yx}, n_{10}^{yx}, n_{11}^{yx})}{\partial n_{11}^{yx}} \right|_{\mu} = \left. -\frac{n_{01}^{yx} n_{10}^{yx}}{(n_{11}^{yx})^2} \right|_{\mu} = -\frac{E(n_{10}^{yx}) E(n_{01}^{yx})}{[E(n_{11}^{yx})]^2} \\ & = -\frac{r_1^{yx} (1 - r_2^{yx}) r_2^{yx} (1 - r_1^{yx})}{(r_1^{yx} r_2^{yx})^2} = -\frac{(1 - r_2^{yx})(1 - r_1^{yx})}{r_1^{yx} r_2^{yx}}. \end{aligned}$$

Par suite,

$$\begin{aligned}
V\left(\widehat{n_{00}^{yx}}\right) &\simeq n^{yx} \left[\left(\frac{1-r_2^{yx}}{r_2^{yx}} \right)^2 (1-r_1^{yx}) r_2^{yx} (1-(1-r_1^{yx}) r_2^{yx}) \right. \\
&\quad + \left(\frac{1-r_1^{yx}}{r_1^{yx}} \right)^2 r_1^{yx} (1-r_2^{yx}) (1-r_1^{yx} (1-r_2^{yx})) \\
&\quad \left. + \left(\frac{(1-r_2^{yx})(1-r_1^{yx})}{r_1^{yx} r_2^{yx}} \right)^2 r_1^{yx} r_2^{yx} (1-r_1^{yx} r_2^{yx}) \right]
\end{aligned}$$

En référence aux distributions marginales et les covariance respectivement définies a (1.52) et (??), on obtient :

$$\begin{aligned}
V\left(\widehat{n_{00}^{yx}}\right) &\simeq n^{yx} \left[(1-r_2^{yx})^2 (1-r_1^{yx}) \left(\frac{1}{r_2^{yx}} - (1-r_1^{yx}) \right) \right. \\
&\quad + (1-r_1^{yx})^2 (1-r_2^{yx}) \left(\frac{1}{r_1^{yx}} - (1-r_2^{yx}) \right) \\
&\quad \left. + (1-r_2^{yx})^2 (1-r_1^{yx})^2 \left(\frac{1}{r_1^{yx} r_2^{yx}} - 1 \right) \right] \\
&\simeq n^{yx} (1-r_2^{yx}) (1-r_1^{yx}) \left\{ (1-r_2^{yx}) \left[\frac{1}{r_2^{yx}} - (1-r_1^{yx}) \right] \right. \\
&\quad \left. + (1-r_1^{yx}) \left[\frac{1}{r_1^{yx}} - (1-r_2^{yx}) \right] + (1-r_2^{yx}) (1-r_1^{yx}) \left(\frac{1}{r_1^{yx} r_2^{yx}} - 1 \right) \right\} \\
&\simeq n^{yx} (1-r_2^{yx}) (1-r_1^{yx}) \left[\frac{1-r_2^{yx}}{r_2^{yx}} + \frac{1-r_1^{yx}}{r_1^{yx}} \right. \\
&\quad \left. + (1-r_1^{yx}) (1-r_2^{yx}) \left(\frac{1}{r_1^{yx} r_2^{yx}} - 3 \right) \right]
\end{aligned}$$

Les équations (??) et (1.30) permettent d'écrire :

$$\begin{aligned}
V\left(\widehat{n_{00}^{yx}}\right) &\simeq n^{yx} \left(\frac{n_{01}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) \left(\frac{n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right) \left\{ \frac{\frac{n_{01}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}}{\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}} + \frac{\frac{n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}}{\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}} \right. \\
&\quad \left. + \left(\frac{n_{01}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) \left(\frac{n_{10}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right) \left[\frac{1}{\left(\frac{n_{11}^{yx}}{n_{01}^{yx} + n_{11}^{yx}}\right) \left(\frac{n_{11}^{yx}}{n_{10}^{yx} + n_{11}^{yx}}\right)} - 3 \right] \right\} \\
&\simeq n^{yx} \left[\frac{n_{01}^{yx} n_{10}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right] \left\{ \frac{n_{01}^{yx}}{n_{11}^{yx}} + \frac{n_{10}^{yx}}{n_{11}^{yx}} \right. \\
&\quad \left. + \left[\frac{n_{01}^{yx} n_{10}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \right] \left[\frac{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx}) - 3(n_{11}^{yx})^2}{(n_{11}^{yx})^2} \right] \right\} \\
&\simeq n^{yx} \left(\frac{\widehat{n_{00}^{yx}}}{n^{yx}}\right) \left[\frac{n_{01}^{yx}}{n_{11}^{yx}} + \frac{n_{10}^{yx}}{n_{11}^{yx}} + \left(\frac{\widehat{n_{00}^{yx}}}{n^{yx}}\right) \left(\frac{n^{yx}}{n_{11}^{yx}} - 3\right) \right]
\end{aligned}$$

et donc,

$$\begin{aligned}
V\left(\widehat{n_{00}^{yx}}\right) &\simeq \widehat{n_{00}^{yx}} \left[\frac{n_{01}^{yx}}{n_{11}^{yx}} + \frac{n_{10}^{yx}}{n_{11}^{yx}} + \widehat{n_{00}^{yx}} \left(\frac{1}{n_{11}^{yx}} - \frac{3}{n^{yx}} \right) \right] \\
&\simeq \widehat{n_{00}^{yx}} \left[\frac{1}{n_{11}^{yx}} \left(n_{10}^{yx} + n_{01}^{yx} + \widehat{n_{00}^{yx}} \right) - \frac{3\widehat{n_{00}^{yx}}}{n^{yx}} \right].
\end{aligned} \tag{1.66}$$

1.A.7.2 Expression de \widehat{p} et sa variance

Expression de \widehat{p} . En remarquant que

$$\widehat{p_{00}^{yx}} = \left(1 - \widehat{r_1^{yx}}\right) \left(1 - \widehat{r_2^{yx}}\right) = \frac{n_{01}^{yx} n_{10}^{yx}}{(n_{01}^{yx} + n_{11}^{yx})(n_{10}^{yx} + n_{11}^{yx})} \tag{1.67}$$

et en utilisant l'expression de $\widehat{n^{yx}}$ donnée à l'équation (1.13), l'expression de l'estimateur \widehat{p} est donnée par :

$$\begin{aligned}
\widehat{p} &= \frac{\widehat{n}^{11}\widehat{p}_{00}^{11} + \widehat{n}^{01}\widehat{p}_{00}^{01}}{\widehat{n}^{11} + \widehat{n}^{01}} \\
&= \frac{\left[\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}} \right] \left[\frac{n_{01}^{11}n_{10}^{11}}{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})} \right]}{\left[\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}} \right] + \left[\frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}} \right]} \\
&\quad + \frac{\left[\frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}} \right] \left[\frac{n_{01}^{01}n_{10}^{01}}{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})} \right]}{\left[\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}} \right] + \left[\frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}} \right]} \\
&= \frac{\frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}} + \frac{n_{01}^{01}n_{10}^{01}}{n_{11}^{01}}}{\frac{n_{01}^{11}(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + n_{11}^{11}(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}n_{11}^{11}}} \\
&= \frac{n_{01}^{11}n_{10}^{11}n_{11}^{01} + n_{01}^{01}n_{10}^{01}n_{11}^{11}}{n_{11}^{01}(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + n_{11}^{11}(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}. \tag{1.68}
\end{aligned}$$

La variance de \widehat{p} est telle que

$$V\left(\widehat{p}_{00}^{11} - \widehat{p}_{00}^{01}\right) = \widehat{p}\left(1 - \widehat{p}\right)\left[\left(\widehat{n}^{11}\right)^{-1} + \left(\widehat{n}^{01}\right)^{-1}\right].$$

On peut remarquer que

$$\begin{aligned}
\frac{1}{\widehat{n}^{11}} + \frac{1}{\widehat{n}^{01}} &= \frac{\widehat{p}_{00}^{11}}{\widehat{n}_{00}^{11}} + \frac{\widehat{p}_{00}^{01}}{\widehat{n}_{00}^{01}} \\
&= \frac{\frac{n_{01}^{11}n_{10}^{11}}{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}}{\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}}} + \frac{\frac{n_{01}^{01}n_{10}^{01}}{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}}{\frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}}} \\
&= \frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}} + \frac{n_{01}^{01}n_{10}^{01}}{n_{11}^{01}}
\end{aligned}$$

et

$$\begin{aligned}
\widehat{p} \left(1 - \widehat{p} \right) &= \left[\frac{n_{01}^{11} n_{10}^{11} n_{11}^{01} + n_{01}^{01} n_{10}^{01} n_{11}^{11}}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right] \\
&\times \left[1 - \frac{n_{01}^{11} n_{10}^{11} n_{11}^{01} + n_{01}^{01} n_{10}^{01} n_{11}^{11}}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right] \\
&= \frac{n_{01}^{11} n_{10}^{11} n_{11}^{01} + n_{01}^{01} n_{10}^{01} n_{11}^{11}}{[n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})]} \\
&\times \left[\frac{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right. \\
&\quad \left. - \frac{n_{01}^{11} n_{10}^{11} n_{11}^{01} + n_{01}^{01} n_{10}^{01} n_{11}^{11}}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right] \\
&= \left[\frac{n_{01}^{11} n_{10}^{11} n_{11}^{01} + n_{01}^{01} n_{10}^{01} n_{11}^{11}}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right] \\
&\times \left[\frac{n_{11}^{01} \left(n_{01}^{11} n_{11}^{11} + n_{10}^{11} n_{11}^{11} + (n_{11}^{11})^2 \right) + n_{11}^{11} \left(n_{01}^{01} n_{11}^{01} + n_{10}^{01} n_{11}^{01} + (n_{11}^{01})^2 \right)}{n_{11}^{01} (n_{01}^{11} + n_{11}^{11}) (n_{10}^{11} + n_{11}^{11}) + n_{11}^{11} (n_{01}^{01} + n_{11}^{01}) (n_{10}^{01} + n_{11}^{01})} \right] \\
&= \left[\frac{\frac{n_{01}^{11} n_{10}^{11}}{n_{11}^{11}} + \frac{n_{01}^{01} n_{10}^{01}}{n_{11}^{01}}}{\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}} + \frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}}} \right] \\
&\times \left[\frac{(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})}{\frac{(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11})}{n_{11}^{11}} + \frac{(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})}{n_{11}^{01}}} \right].
\end{aligned}$$

Par suite,

$$\begin{aligned}
V\left(\widehat{p_{00}^{11}} - \widehat{p_{00}^{01}}\right) &= \left[\frac{\frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}} + \frac{n_{01}^{01}n_{10}^{01}}{n_{11}^{01}}}{\frac{(n_{01}^{11}+n_{11}^{11})(n_{10}^{11}+n_{11}^{11})}{n_{11}^{11}} + \frac{(n_{01}^{01}+n_{11}^{01})(n_{10}^{01}+n_{11}^{01})}{n_{11}^{01}}} \right] \\
&\times \left[\frac{(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})}{\frac{(n_{01}^{11}+n_{11}^{11})(n_{10}^{11}+n_{11}^{11})}{n_{11}^{11}} + \frac{(n_{01}^{01}+n_{11}^{01})(n_{10}^{01}+n_{11}^{01})}{n_{11}^{01}}} \right] \\
&\times \left[\frac{n_{01}^{11}n_{10}^{11}}{n_{11}^{11}} + \frac{n_{01}^{01}n_{10}^{01}}{n_{11}^{01}} \right] \\
&= \left[\frac{n_{11}^{01}n_{01}^{11}n_{10}^{11} + n_{11}^{11}n_{01}^{01}n_{10}^{01}}{n_{11}^{01}(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + n_{11}^{11}(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})} \right] \\
&\times \left[\frac{n_{11}^{01}n_{11}^{11}[(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})]}{n_{11}^{01}(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + n_{11}^{11}(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})} \right] \\
&\times \left[\frac{n_{11}^{01}n_{01}^{11}n_{10}^{11} + n_{11}^{11}n_{01}^{01}n_{10}^{01}}{n_{11}^{01}n_{11}^{11}} \right] \\
&= \frac{(n_{11}^{01}n_{01}^{11}n_{10}^{11} + n_{11}^{11}n_{01}^{01}n_{10}^{01})^2 [(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})]}{[n_{11}^{01}(n_{01}^{11} + n_{11}^{11})(n_{10}^{11} + n_{11}^{11}) + n_{11}^{11}(n_{01}^{01} + n_{11}^{01})(n_{10}^{01} + n_{11}^{01})]^2} \quad (1.69) \\
&= \left(\widehat{p_{00}^{11}} + \widehat{p_{00}^{01}}\right)^2 [(n_{01}^{11} + n_{10}^{11} + n_{11}^{11}) + (n_{01}^{01} + n_{10}^{01} + n_{11}^{01})]
\end{aligned}$$

1.A.7.3 Variance et covariance des estimateurs $\widehat{\beta}_0$ et $\widehat{\beta}_1$

Variance de $\widehat{\beta}_0$. D'après le système établi en (1.5), l'estimateur $\widehat{\beta}_0$ est tel que

$$\widehat{\Phi}(\widehat{\beta}_0) \equiv \Phi(\widehat{\beta}_0) = \widehat{q}.$$

En développant $\Phi(\widehat{\beta}_0)$ à l'ordre un au voisinage de β_0 , on a :

$$\begin{aligned}
\Phi(\widehat{\beta}_0) &\simeq \Phi(\beta_0) + (\widehat{\beta}_0 - \beta_0) \Phi'(\beta_0) \\
\Rightarrow V(\Phi(\widehat{\beta}_0)) &\simeq (\Phi'(\beta_0))^2 V(\widehat{\beta}_0).
\end{aligned}$$

Ainsi,

$$\begin{aligned}
V(\Phi(\widehat{\beta}_0)) &= V(\widehat{q}) \Rightarrow (\Phi'(\beta_0))^2 V(\widehat{\beta}_0) \simeq V(\widehat{q}) \\
\Rightarrow V(\widehat{\beta}_0) &\simeq (\Phi'(\beta_0))^{-2} V(\widehat{q}).
\end{aligned}$$

Variance de $\widehat{\beta}_1$. De même,

$$\Phi(\widehat{\beta_0 + \beta_1}) \equiv \Phi(\widehat{\beta_0} + \widehat{\beta_1}) = \widehat{q}.$$

Or,

$$\Phi(\widehat{\beta_0} + \widehat{\beta_1}) \simeq \Phi(\beta_0 + \beta_1) + (\widehat{\beta_0} - \beta_0) \Phi'(\beta_0) + (\widehat{\beta_1} - \beta_1) \Phi'(\beta_1)$$

et donc

$$\begin{aligned} V(\Phi(\widehat{\beta_0} + \widehat{\beta_1})) &\simeq (\Phi'(\beta_0))^2 V(\widehat{\beta_0}) + (\Phi'(\beta_1))^2 V(\widehat{\beta_1}) \\ &\quad + 2\Phi'(\beta_0) \Phi'(\beta_1) Cov(\widehat{\beta_0}, \widehat{\beta_1}). \end{aligned}$$

Mais

$$\begin{aligned} Cov(\widehat{\beta_0}, \widehat{\beta_1}) &= E[(\widehat{\beta_0} - \beta_0)(\widehat{\beta_1} - \beta_1)] \\ &= E\{[\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))] \\ &\quad \times [\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q})) - \Phi^{-1}(E(\widehat{q})) + \Phi^{-1}(E(\widehat{q}))]\} \\ &= E\{[\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))][\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))]\} \\ &\quad - E[\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))]^2 \\ &= -V(\widehat{\beta_0}) \end{aligned}$$

car, sous l'hypothèse d'indépendance

$$E\{[\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))][\Phi^{-1}(\widehat{q}) - \Phi^{-1}(E(\widehat{q}))]\} = 0$$

Ainsi,

$$\begin{aligned} V(\Phi(\widehat{\beta_0} + \widehat{\beta_1})) &\simeq (\Phi'(\beta_0))^2 V(\widehat{\beta_0}) + (\Phi'(\beta_1))^2 V(\widehat{\beta_1}) \\ &\quad - 2\Phi'(\beta_0) \Phi'(\beta_1) V(\widehat{\beta_0}) \\ &\simeq [\Phi'(\beta_0) - 2\Phi'(\beta_1)] \Phi'(\beta_0) V(\widehat{\beta_0}) + (\Phi'(\beta_1))^2 V(\widehat{\beta_1}). \end{aligned}$$

Enfin,

$$V(\widehat{q}) \simeq [\Phi'(\beta_0) - 2\Phi'(\beta_1)](\Phi'(\beta_0))^{-1} V(\widehat{q}) + (\Phi'(\beta_1))^2 V(\widehat{\beta}_1)$$

ce qui donne

$$V(\widehat{\beta}_1) \simeq \frac{V(\widehat{q}) - [\Phi'(\beta_0) - 2\Phi'(\beta_1)](\Phi'(\beta_0))^{-1} V(\widehat{q})}{(\Phi'(\beta_1))^2}$$

Essai # 2

Arrosage privé à petite échelle : un
modèle d'assurance pour les petites
exploitations agricoles

Résumé. Dans le cadre de ce chapitre, on développe un modèle pour évaluer les bienfaits et les limites de l'arrosage manuel, sur la base de l'utilité et du revenu espéré des exploitants. La sécheresse et l'inondation constituent les deux principaux risques climatiques contre lesquels les petits exploitants doivent se prémunir en faisant des investissements additionnels. Le modèle évalue la possibilité de réduire les pertes pouvant résulter d'un épisode de sécheresse et/ou d'inondation, étant donné l'absence de systèmes automatiques d'arrosage et/ou d'irrigation. Les deux risques climatiques pris en compte - la sécheresse et l'inondation - donnent lieu à quatre « états de la nature » possibles : un épisode de sécheresse suivi par une inondation dévastatrice, un épisode de sécheresse qui n'est pas suivi d'inondation, une inondation sans qu'il y ait eu de sécheresse précédemment et une situation naturelle « clémente » - sans sécheresse ni inondation.

L'arrosage est rentable et augmente le revenu espéré si : a) l'intensité des sécheresses est relativement élevée mais le risque d'inondation est peu important ; b) la corrélation entre risque de sécheresse et risque d'inondation est faible ; c) les éventuelles compensations mises en place par l'autorité publique en cas d'inondation tiennent compte de l'effort spécifique des exploitants. On montre aussi que la garantie d'une aide financière (exogène) et une couverture d'assurance endogène en cas d'inondation ont toutes deux un impact positif sur l'effort d'arrosage. On établit en particulier que *i*) l'effort d'arrosage et le revenu espéré diminuent avec la corrélation (positive) entre les risques d'inondation et de sécheresse, et que, puisque assurance contre le risque d'inondations et efforts d'arrosage sont complémentaires, *ii*) la couverture d'assurance privée (endogène) contre le risque d'inondation est corrélée positivement avec l'effort déployé par les exploitants.

Deux résultats essentiels sont alors établis. Dans un premier temps, on établit qu'il n'est pas possible de trouver un niveau d'effort θ^* et une police d'assurance inondation (privée) A^* tels que le couple (θ^*, A^*) qui maximise l'utilité espérée de l'exploitant soit associé à une situation sans risque pour l'exploitant. En d'autres termes, sauf si les deux risques (de sécheresse et d'inondation) sont totalement indépendants, deux instruments (effort investi et police d'assurance inondation) ne peuvent suffire à offrir une couverture d'assurance totale. Dans un second temps, on montre que, dans certaines situations, il est

impossible d'avoir une assurance inondation rentable et l'exploitant ne fait jamais l'effort d'arroser. Ces situations sont identifiées comme correspondant à des lieux où le risque d'être confronté à une inondation est d'autant plus important qu'on vient de traverser une longue période de sécheresse. Dans de telle circonstances, *ne rien faire* pour ce prémunir des variations climatiques est parfaitement rationnel. Ce chapitre fournit donc un argument théorique pour expliquer les faibles niveaux d'assurance observés dans le secteur agricole.

2.1 Introduction

Les changements climatiques (CC) observés durant ces dernières années se manifestent sous différentes formes telles des sécheresses extrêmes, les changements dans les régimes de précipitation, les inondations, la hausse de la température globale, etc. Les impacts enregistrés (Kilroy, 2015; Rai et Rai, 2013), tant sur le plan biophysique (hydrologique, climatique, écologique, agricole) que socioéconomique (alimentation, prix et production agricoles, nutrition et santé, etc.) de ces changements climatiques ont des conséquences énormes sur les conditions de vie des gens en général à travers le monde et dans les pays en développement (PED) en particulier (Nath et Behera, 2011). Dans un rapport rédigé pour le compte du Parlement Européen en 2007, les impacts des CC répertoriés (Ludwig et al., 2007) sur les PED sont dévastateurs. Augmentation de la température globale, changements dans les régimes de précipitations et conséquemment dans la fréquence et l'intensité de la sécheresse et des inondations. Malgré les progrès remarquables de la technologie -telles par exemple les variétés à rendements plus élevés (Feder et Zilberman, 1985)- destinée à améliorer la viabilité et la productivité du secteur agricole, ce secteur reste le plus dépendant des conditions climatiques (Mendelsohn et Dinar, 2009). Les résultats des projections réalisées ne sont pas encourageants. Selon le rapport du Groupe d'Experts Intergouvernemental sur l'Evolution du Climat (GIEC), sur la base de tous les scénarios d'émissions considérés, les projections indiquent une augmentation de la température de surface au cours du XXIe siècle. Il est très probable que la fréquence et la durée des vagues de chaleur augmenteront et que les précipitations extrêmes vont devenir plus intenses et plus fréquentes dans de nombreuses régions (GIEC et Pachauri, 2014).

Etant donné la forte dépendance de la majorité de la population au secteur agricole dans les PED, la situation décrite précédemment a des conséquences encore beaucoup plus dramatiques et en particulier sur la situation des personnes engagées dans une agriculture de subsistance dans ces pays (Nath et Behera, 2011). Les stratégies adoptées par ces agriculteurs pour faire face à ces risques multiples (risques de production, risques de ressources, risques de récoltes, etc..) sont regroupées en deux grandes catégories (Hazell,

1992) : les stratégies de réduction des risques (diversification des sources de revenu, diversification des cultures, banque de sémences, support d'organisations communautaires, ..) et les stratégies de gestion ex-post des risques (liquidation d'actifs, usage de stocks de nourriture,...). Ces stratégies sont non seulement coûteuses et peu efficaces mais encore se révèlent non effectives dans le cas des risques covariants. Dans ce contexte, des outils de gestion de risque tels des produits d'assurance sur les récoltes et des programmes d'aide publique aux victimes des conséquences des CC sont mis en place. Ces produits d'assurance peuvent contribuer non seulement à protéger les agriculteurs mais encore à mieux les adapter aux CC. Bien que ces produits sont parfois utilisés comme des instruments d'assistance en cas de catastrophe, ils ont aussi leurs désavantages. En particulier, ils peuvent décourager le développement d'autres solutions privées sans pouvoir remplacer complètement une assistance ex-post (Anton et al., 2013b).

De part l'importance du secteur agricole dans ces pays, divers éléments de solution sont proposés tels le rôle important à jouer par les gouvernements dans les PED, la place à donner aux secteurs de la population les plus impliqués dans le secteur agricole, la mise en place de synergie entre les politiques climatiques et d'autres politiques de développement, etc. (Nath et Behera, 2011). Cependant, les personnes engagées dans une agriculture de subsistance ne sont pas en général, touchées par les quelques rares mesures, quand elles existent, de support et d'aide aux agriculteurs impactés par ces conditions climatiques extrêmes.

On se propose dans ce chapitre de thèse de développer un modèle devant évaluer les bienfaits et limites de l'arrosage privé à petite échelle ¹ dans une perspective de réduction des pertes potentielles pouvant résulter d'un épisode sécheresse, étant donné le non accès de ces personnes à un système d'arrosage pour des raisons de la non disponibilité, en général, d'un tel système. L'impact sur l'effort d'arrosage d'une couverture d'assurance sur les inondations sera aussi évalué.

Le chapitre est organisé de la manière suivante. Dans un premier temps, on présente

1. L'arrosage privé à petite échelle ou petite irrigation se définit comme l'apport d'eau sur de petites parcelles contrôlé par les producteurs eux-mêmes, et à l'aide de techniques que les agriculteurs peuvent facilement maîtriser et entretenir (Tillie et al., 2019; Carter, 1989).

les principaux résultats du modèle. On précise, dans un deuxième temps, la nature du modèle et les principales hypothèses simplificatrices utilisées. Les principaux états de la nature auxquels fait face le petit agriculteur sont précisés ainsi que les probabilités et le revenu associés. Par suite, la fonction d'utilité indirecte de l'agriculteur et son niveau de revenu espéré sont définis. Dans un troisième temps, on analyse dans le cadre du modèle, les conséquences de la corrélation ou non des risques de sécheresse et d'inondation sur l'effort d'arrosage et le revenu espéré de l'agriculteur. En quatrième lieu, on apprécie en cas d'inondation l'impact d'une aide exogène et d'une couverture d'assurance endogène sur l'effort d'arrosage et le revenu espéré de l'exploitant. Les principaux résultats du modèle sont ensuite illustrés sur la base d'une fonction d'utilité appropriée dans un cinquième temps. Enfin, on synthétise les principaux résultats du modèle et on précise ses principales limites.

2.2 Les principaux résultats du modèle

Le modèle décrit et analyse, dans un contexte de changement climatique, le comportement des petits exploitants agricoles qui sont supposés faire face à deux risques essentiels : un risque de sécheresse et un risque d'inondation. Ces petits agriculteurs sont supposés être les principaux agents concernés dans cette situation et disposant d'une technologie de production qu'ils maîtrisent. La sécheresse et l'inondation constituent les deux principaux risques climatiques contre lesquels ils doivent se prémunir en faisant des investissements additionnels. Le gouvernement est supposé intervenir, quand il le fait, à posteriori dans le cas des épisodes de sécheresse ou d'inondation. Par ailleurs, les entreprises d'assurance sont supposées offrir des produits partiellement ou non adaptés aux conditions socioéconomiques des petits exploitants et à la mitigation des risques climatiques et difficilement accessibles.

Dans ces conditions, les principaux résultats du modèle peuvent se résumer comme suit :

1. Un effort d'arrosage est faisable et bénéfique dans le cas où les risques de sécheresse

et d'inondation sont négativement corrélés.

2. L'existence de risques inhérents rend impossible une couverture d'assurance complète
3. Il existe un couple optimal (θ^*, A^*) qui maximise l'utilité espérée de l'exploitant où θ^* =niveau d'effort et A^* =couverture d'assurance
4. En présence de risques corrélés de sécheresse et d'inondation, l'impossibilité d'avoir une assurance inondation rentable et où l'exploitant ne fait jamais l'effort d'arroser, sont des situations sans espoir. Ces situations correspondent naturellement à celles où il est de plus en plus probable qu'il survienne une inondation, quand il apparaît de moins en moins probable qu'il s'installe une sécheresse, décourageant ainsi le moindre effort d'arrosage.

2.3 Le modèle

Le modèle est monopériodique en ce sens que l'on suppose que tout se passe dans un espace de temps relativement réduit - une année par exemple. Les petits agriculteurs sont les principaux agents concernés dans cette situation. Ils sont supposés disposer d'une technologie de production qu'ils maîtrisent. La sécheresse et l'inondation constituent les deux principaux risques climatiques contre lesquels ils doivent se prémunir en faisant des investissements additionnels. Le gouvernement est supposé intervenir, quand il le fait, à posteriori dans le cas des épisodes de sécherese ou d'inondation. C'est la situation qui s'observe le plus souvent dans beaucoup de pays à faible revenu (Barrett, 2011). Les entreprises d'assurance sont supposées offrir des produits partiellement ou non adaptés aux conditions socioéconomiques des petits exploitants, à la mitigation des risques climatiques et difficilement accessibles. Tout cela est dû en partie au faible niveau de développement des pays à faible revenu, à des contraintes du côté de l'offre et de la demande des produits d'assurance, en particulier les produits de couverture de risques climatiques telles que soulignées par CII - The Chartered Insurance Institute - (CII, 2009).

Les personnes engagées dans une agriculture de subsistance font face essentiellement

à deux risques d'ordre climatique : la sécheresse, qui peut survenir avec une probabilité p ($0 < p < 1$) et son contraire, une sur-abondance de pluies menant à inondation (et la destruction totale des récoltes), avec une probabilité d'occurrence q ($0 < q < 1$). Contrairement aux pluies torrentielles et aux cyclones, contre lesquels il est difficile de se prémunir, il est possible de lutter contre la sécheresse en arrosant.

Supposons dans un premier temps qu'un petit agriculteur, ayant un niveau de *revenu initial* R , décide d'adopter une stratégie d'arrosage pour se protéger contre la sécheresse. Étant donné le risque d'inondation, le petit exploitant fait face à quatre situations possibles. Il peut faire face à une situation de sécheresse (S) suivie d'un épisode d'inondation (I) avec une probabilité $\pi_{SI} = P(S \cap I)$ telle que $\pi_{SI} < p, q$; il peut tout aussi, avec une probabilité $\pi_{\emptyset I}$, en absence de sécheresse, subir les conséquences d'un épisode d'inondation ou tout au contraire, en absence d'inondation, faire face à la sécheresse avec une probabilité $\pi_{S\emptyset}$; Enfin, l'exploitant peut bénéficier des conditions climatiques clémentes avec une probabilité $\pi_{\emptyset\emptyset}$. L'exploitant fait donc face à un ensemble de situations

$$\Omega = \{SI, \emptyset I, S\emptyset, \emptyset\emptyset\} \quad (2.1)$$

avec une probabilité $P(w_j) = \pi_j$, $w_j \in \Omega$ telle que $0 < \pi_j < p, q < 1$. La situation est résumée dans le tableau suivant, Π étant le profit réalisé :

TABLE 2.3.1 – Les différents états de la nature et les probabilités associées

États \longrightarrow	SI	$\emptyset I$	$S\emptyset$	$\emptyset\emptyset$
Probabilité	π_{SI}	$\pi_{\emptyset I}$	$\pi_{S\emptyset}$	$\pi_{\emptyset\emptyset}$
	π_{SI}	$q - \pi_{SI}$	$p - \pi_{SI}$	$1 - (p + q - \pi_{SI})$
Revenu sans arrosage	R	R	$R + (1 - \alpha)\Pi$	$R + \Pi$
Revenu avec arrosage	$R - c\theta$	R	$R - c\theta + [\alpha\theta + (1 - \alpha)]\Pi$	$R + \Pi$

Les probabilités p et q sont telles que $p + q \geq 1$ du fait qu'une situation de sécheresse intense peut être suivie par un épisode d'inondation dévastatrice. Le paramètre $\theta \in [0; 1]$ indique l'effort d'arrosage consenti par l'exploitant et $\alpha \in [0; 1]$ dénote le degré de sévérité de la sécheresse. On s'attend à ce que plus α est petit (la sécheresse est moins sévère), plus θ sera grand. On suppose que l'effort d'arrosage θ est coûteux et que les coûts d'arrosage

liés à l'intensité de l'effort déployé sont linéaires et égaux à $c\theta^2$.

On suppose que l'exploitant est averse au risque et que sa fonction d'utilité $v(\cdot)$ est deux fois dérivable, croissante et concave c'est-à-dire $v'(\cdot) > 0$ et $v''(\cdot) < 0$. L'espérance de sa fonction d'utilité (indirecte), dans le cas où *il fait un effort d'arrosage* θ , est définie par :

$$\begin{aligned} E(V) &= \pi_{SI}v(R - c\theta) + (q - \pi_{SI})v(R) \\ &+ (p - \pi_{SI})v(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &+ (1 - p - q + \pi_{SI})v(R + \Pi). \end{aligned} \quad (2.2)$$

Son revenu espéré $E(R)$ correspondant est :

$$E(R) = R + (1 - q)\Pi + (p - \pi_{SI})(\theta - 1)\alpha\Pi - pc\theta \quad (2.3)$$

2. En effet, si on considère une fonction de coût plus générale de classe \mathcal{C}^1 , $c(\theta)$, $\theta \in [0; 1]$, croissante ($c'(\theta) > 0$) et convexe ($\forall a, b \in]0, 1[\text{ tq } a < b, \text{ on a : } c'(a) \leq \frac{c(b) - c(a)}{b - a} \leq c'(b)$), les résultats généraux du modèle ne changent pas. Le revenu net de l'exploitant, lui, peut changer dépendamment de la structure des coûts.

De même, les résultats généraux du modèle ne changent pas si on considère des structures de coûts différents en fonction des régions. Dans ce cas, on peut définir des fonctions de coût régionales $c_r(\theta)$ ayant les propriétés précédemment explicitées. La fonction de coût total $c(\theta)$ s'écrit alors : $c(\theta) = \sum_{r=1}^R c_r(\theta)$.

Encadré 2.3. Haïti : La situation des exploitations agricoles

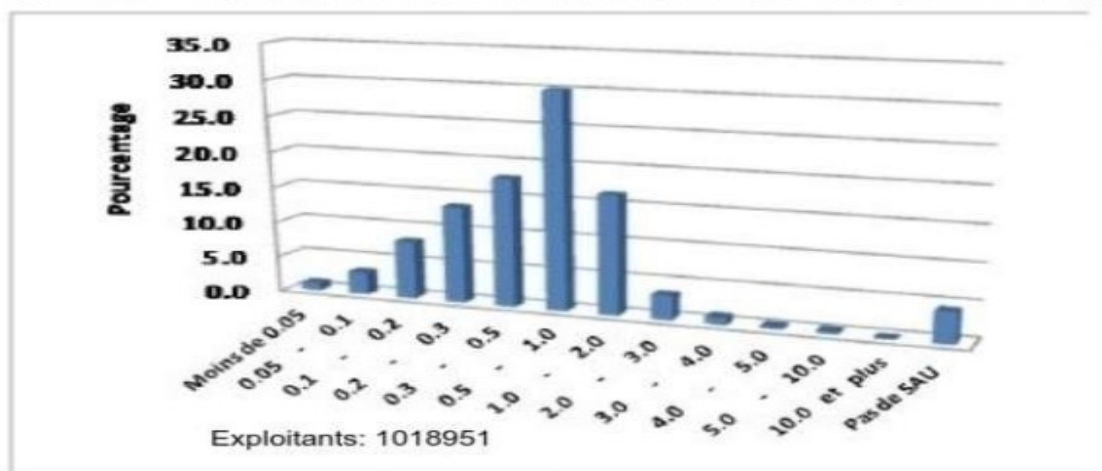
Du point de vue des catastrophes naturelles, Haïti, qui fait face à un taux élevé de déforestation couplé à des pluies intenses, a été frappée durant les trente dernières années par six (6) ouragans. Les départements de l'Ouest et du Sud ont été les plus affectés du fait qu'ils se situent dans la trajectoire des ouragans violents. Port-au-Prince, la capitale, avec une proportion importante de ses habitants vivant dans des plaines inondables et dans des logements inadaptés, est particulièrement vulnérable aux inondations. Les départements du Nord-ouest, de l'Artibonite, du Nord-est et du Centre sont fréquemment affectés par une sécheresse récurrente découlant d'un schéma pluvieux erratique couplé à une infrastructure limitée de gestion de l'eau.

Dans ce contexte précédemment décrit, le secteur agricole haïtien est grandement affecté. En particulier, la situation des petits exploitants, majoritaires dans le secteur, est catastrophique. Les données les plus représentatives et les plus complètes disponibles à date, bien que peu récentes, proviennent du Recensement Général Agricole (RGA) de 2009 réalisé par le Ministère de l'Agriculture, des ressources Naturelles et du Développement Rural (MARNDR). Le tableau (a) et la figure (b) construits à partir de ces données donnent une bonne idée de la configuration des exploitations agricoles. En effet, 73.9% des exploitations agricoles ont une superficie de moins d'un carreau soit autant d'agriculteurs qui exploitent moins d'un carreau. De plus, 98.6% des exploitations sont gérées par les exploitants. Du total des parcelles cultivées, seulement 12.9% sont irriguées.

(a) Caractéristiques des exploitations agricoles

Exploitations agricoles recensées	: 1 018 951 (736 241.5 cr)
Nombre d'Exploitations agricoles de moins d'1 cr	: 73.9% (Petite Taille)
Pourcentage de Surface Agricole Utilisée (SAU) de moins d'1 cr	: 46.7%
Exploitants agricoles recensés	: 1 018 951
Exploitations gérées directement par les exploitants	: 98.6%
Parcelles irriguées	: 234 587 (12.9%)
Total parcelles	: 1 818 504 (87.1%)
Source: RGA de 2009 , MARNDR	Note: 1 cr (carreau) = 1 ha 29

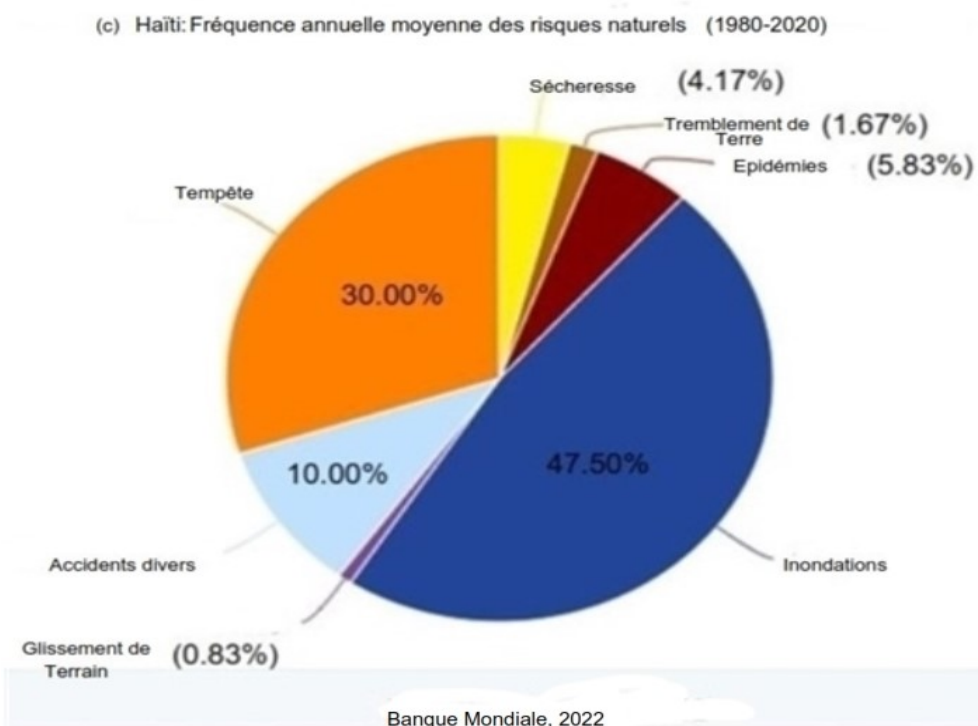
(b) Répartition des exploitants par classe taille de SAU (Surface Agricole Utilisée)



Source: MARNDR, RGA 2009

Encadré 2.3. Haïti : La situation des exploitations agricoles (suite)

Haïti est un pays très exposé et très vulnérable aux catastrophes naturelles. Cette vulnérabilité se voit aggravée par les conditions environnementales critiques et l'existence de logements majoritairement mal construits localisés dans espaces inappropriés. Sur la période 1980-2020, le pays a été frappé majoritairement par des inondations. En effet, celles-ci constituent 47.5% (figure (c)), des événements naturels qu'a connus le pays. Les Tempêtes viennent en deuxième position avec 30% d'occurrence.



La sécheresse constitue seulement 4.17% des événements naturels ayant impacté le pays sur cette période. Il est cependant important de garder à l'esprit que ces données reflètent la fréquence des différents événements climatiques - mais pas leur gravité et encore moins leurs coûts. Plus intéressant, pour mieux apprécier la modélisation faite dans cette thèse est l'observation suivante : l'existence de zones géographiques où des épisodes de sécheresse et d'inondation se succèdent comme le montrent les graphiques de la page suivante relatifs à l'île d'Haïti (la République d'Haïti et la République Dominicaine).

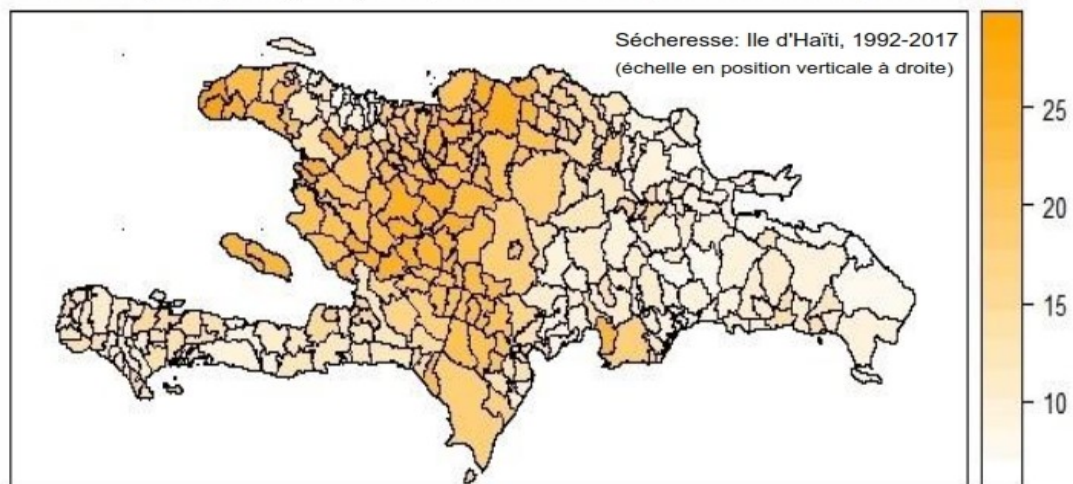
Encadré 2.3. Haïti : La situation des exploitations agricoles (suite)

Sur la période 1992-2017, des croquis cartographiques disponibles permettent de visualiser les différentes zones géographiques de l'île affectées par des épisodes de sécheresse et d'inondation (figures (d) et (e)). Située à l'Est de l'île, Haïti a connu sur la période un grand nombre d'épisodes de sécheresse (figure (d)) à intensité plus ou moins variable.

(d) Sécheresse : Ile d'Haïti, 1992-2017



Nombre moyen de jours (0 - 30) sans pluie sur une année pendant le mois le plus sec



Les départements du Sud, du Sud-Est et de la Grande Anse ont été globalement épargnés. De plus, il est remarquable d'observer l'occurrence des épisodes de sécheresse et d'inondation qui se succèdent comment le montrent les graphiques de la page suivante.

Encadré 2.3. Haïti : La situation des exploitations agricoles (suite)

Sur la période 1992-2017, les épisodes d'inondation (figure (e)) observés n'ont pas été si nombreux comparés aux épisodes de sécheresse enregistrés (figure (d)). En effet, les épisodes d'inondation observés sont localisés à Port-de-Paix et Mole Saint Nicolas (département du Nord'Ouest), à l'île de Gonave (département de l'Ouest), à Belle Anse (département du Sud-Est) et aux Cayes (département du Sud).

(e) Inondations : île d'Haïti, 1992-2017



Il est remarquable d'observer l'occurrence des épisodes de sécheresse et d'inondation qui se succèdent notamment pour Port-de-Paix et Mole Saint Nicolas (département du Nord'Ouest), l'île de Gonave (département de l'Ouest) et Belle Anse (département du Sud-Est).

2.3.1 Corrélation entre les risques de sécheresse et d'inondation et effort d'arrosage

On suppose que, comme caractérisée précédemment, la survenue d'une séquence de sécheresse S avec une probabilité p et/ou d'un épisode d'inondation I avec une probabilité q constituent deux états de la nature auxquels doit faire l'agriculteur. L'effort d'arrosage à fournir pour faire face à ces états de la nature est aussi fonction du degré de corrélation des risques d'inondation $R_S = \{(s, p, x_s)\}$ et de sécheresse $R_I = \{(i, q, x_i)\}$. Les variables s et i désignent respectivement les dommages observés pour S et I tels que $p = P(S = s)$ et $q = P(I = i)$; x_s et x_i étant les mesures respectives des dommages enregistrés pour S et I . Pour évaluer l'impact de la corrélation sécheresse-inondation sur l'effort d'arrosage, on commence par déterminer le coefficient de corrélation linéaire entre S et I .

Soit σ_S et σ_I les écarts-type respectifs des états S et I de la nature. Le coefficient de corrélation linéaire ρ_{SI} entre ces deux états est donné par³ :

3. La distribution du couple (S, I) correspond à une loi bivariable de Bernoulli. Sa fonction de masse conjointe p_{SI} définie dans $\{0, 1\} \times \{0, 1\}$ a pour expression (Dai et Wahba, 2013) :

$$p_{SI}(s, i) = p_{11}^{si} p_{10}^{s(1-i)} p_{01}^{(1-s)i} p_{00}^{(1-s)(1-i)}, (s, i) \in \{0, 1\}^2$$

où $p_{si} = P(S = s, I = i)$. On note : $(S, I) \sim \mathcal{B}_2(p_{11}, p_{10}, p_{01}, p_{00})$ tel que $p_{11} + p_{10} + p_{01} + p_{00} = 1$. Les distributions marginales de S et I sont des variables de Bernoulli telles que :

$$S \sim \mathcal{B}(p = p_{10} + p_{11}), I \sim \mathcal{B}(q = p_{01} + p_{11})$$

La covariance entre S et I équivaut à :

$$\begin{aligned} Cov(S, I) &\equiv E[(S - p)(I - q)] = E(SI) - E(S)E(I) \\ &= \left[\sum_{s, i \in \{0, 1\}} si p_{SI}(s, i) \right] - pq \\ &= \pi_{SI} - pq, E(SI) = \pi_{SI}, s = i = 1 \\ &= p_{11} - (p_{10} + p_{11})(p_{01} + p_{11}) \\ &= p_{11}p_{00} - p_{10}p_{01} \end{aligned}$$

Par suite, le coefficient de corrélation linéaire ρ_{SI} entre S et I est donné par

$$\begin{aligned} \rho_{SI} &\equiv \frac{Cov(S, I)}{\sigma_S \sigma_I} = \frac{p_{11}p_{00} - p_{10}p_{01}}{[(p_{10} + p_{11})(p_{01} + p_{00})(p_{01} + p_{11})(p_{10} + p_{00})]^{1/2}} \\ &= \frac{\pi_{SI} - pq}{[pq(1-p)(1-q)]^{1/2}}. \end{aligned}$$

$$\begin{aligned}\rho_{SI} &= \frac{Cov(S, I)}{\sigma_S \sigma_I} \\ &= \frac{\pi_{SI} - pq}{[pq(1-p)(1-q)]^{1/2}}.\end{aligned}\tag{2.4}$$

Analysons maintenant l'impact de la corrélation sécheresse-inondation sur l'effort d'arrosage. En considérant l'équation (2.2), la condition du premier ordre (CPO) par rapport au niveau d'effort θ s'exprime, entre autres variables, en fonction des seules probabilité p et π_{SI} :

$$\begin{aligned}\frac{\partial E(V)}{\partial \theta} &= -\pi_{SI} c v'(R - c\theta) \\ &\quad + (p - \pi_{SI}) [\alpha \Pi - c] v'(R + [\alpha \theta + (1 - \alpha)] \Pi - c\theta).\end{aligned}\tag{2.5}$$

Le niveau d'effort optimal pour l'arrosage dépend donc de *la probabilité qu'une sécheresse soit suivie par une inondation dévastatrice* $\pi_{SI} = P(S \cap I) = E(SI)$.

Si à chaque fois qu'il y a une sécheresse, il y a presque toujours aussi une inondation⁴ alors $\pi_{SI} \rightarrow p$. Il existe donc une forte association entre la survenue d'un épisode de sécheresse S suivie par celui d'une inondation I et donc une corrélation ρ_{SI} significativement

4. Cette hypothèse est empiriquement soutenue par différentes recherches réalisées dans un contexte de changement climatique. Utilisant des séries temporelles sur la période 1961-2015, Fang et al. (2019) et Wu et She (2017) ont établi l'existence de niveaux de risques très élevés de transitions brusques entre périodes de sécheresse et périodes fortement pluviales surtout durant les saisons printemps-été et été-automne en Chine, en particulier dans le Bassin de la rivière des Perles. Ces chercheurs ont aussi vérifié empiriquement que des périodes prolongées de sécheresse ou d'inondations sont plus fréquentes dans ce contexte de changement climatique.

élevée ($\rho_{SI} \rightarrow 1$)⁵. Dans ce cas,

$$\frac{\partial E(V)}{\partial \theta} \rightarrow -pcv'(R - c\theta) < 0, \forall \theta \in [0, 1]. \quad (2.6)$$

Par suite, la destruction totale des récoltes résultant de l'épisode d'inondation décourage tout effort d'arrosage car quelque soit l'effort d'arrosage consenti, l'utilité indirecte de l'exploitant est négative et donc $\theta = 0$. Si la survenue d'une sécheresse et celle d'une inondation sont indépendantes alors $\pi_{SI} \rightarrow pq$ et donc $\rho_{SI} \rightarrow 0$. Dans ce cas, pourvu que la part de profit réalisée dépendamment de l'intensité de la sécheresse surpasse le coût marginal de l'effort déployé ($\alpha\Pi > c$), on a :

$$\frac{\partial E(V)}{\partial \theta} \rightarrow p(1 - q)[\alpha\Pi - c]v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) - pqcv'(R - c\theta). \quad (2.7)$$

On peut remarquer que la fonction

$$w(\theta) = p(1 - q)[\alpha\Pi - c]v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) - pqcv'(R - c\theta)$$

est monotone décroissante car $\forall \theta \in [0, 1]$,

$$\frac{dw(\theta)}{d\theta} = p(1 - q)(\alpha\Pi - c)^2 v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) + c^2 pqv''(R - c\theta) < 0.$$

Par conséquent, l'exploitant n'est pas incité à fournir le moindre effort d'arrosage du fait que tout effort d'arrosage réduit son utilité (indirecte). Les petits agriculteurs sont

5. Quand $\pi_{SI} \rightarrow p$ ou q ,

$$\begin{aligned} \rho_{SI} &= \frac{\pi_{SI} - pq}{[pq(1 - p)(1 - q)]^{1/2}} \\ &\rightarrow \frac{q(1 - q)}{[q^2(1 - q)^2]^{1/2}} \rightarrow 1. \end{aligned}$$

Quand $\pi_{SI} \rightarrow pq$, $\rho_{SI} \rightarrow 0$.

Par contre, quand $\pi_{SI} \rightarrow 0$,

$$\rho_{SI} \rightarrow - \left[\frac{pq}{(1 - p)(1 - q)} \right]^{1/2} < 0.$$

donc averses et vulnérables au risque (Gollier, 1996) et aussi averses aux risques positivement corrélés si on s'en tient aux travaux de Epstein et Tanny (1980) et Denuit et Rey (2010). La possibilité de réalisation indépendante de ces deux états de la nature impacte négativement l'utilité de l'exploitant averse au risque et qui fait face à un risque bivarié indépendant. Le résultat est qu'il s'abstient d'investir le moindre effort du fait de l'existence d'une corrélation positive ρ_{SI} même à la limite de zéro, entre les risques de sécheresse et d'inondation, corrélation qui dénote l'attitude des agents averses au risque à investir dans des entreprises exhibant des risques bivariés positivement corrélés (Denuit et Rey, 2010).

Par contre, si les années de sécheresse sont aussi celles où il n'y a pratiquement jamais d'inondation ce qui implique $\pi_{SI} \rightarrow 0$. Il existe donc une association négative entre la survenue d'un épisode de sécheresse S suivie par celui d'une inondation I et donc $\rho_{SI} < 0$. Par suite, à condition que la part de profit réalisée dépendamment de l'intensité de la sécheresse surpasse le coût marginal de l'effort déployé ($\alpha\Pi > c$),

$$\frac{\partial E(V)}{\partial \theta} \rightarrow p(\alpha\Pi - c)v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) > 0. \quad (2.8)$$

Par conséquent, si les années de sécheresse sont aussi celles où il n'y a pratiquement jamais d'inondation ($\rho_{SI} < 0$), l'exploitant est sûr de récupérer les fruits de l'effort d'arrosage. Dans ce cas, il est incité à fournir un effort de plus en plus important pouvant aller jusqu'à un effort maximal $\theta = 1$ car $\forall \theta \in [0, 1], (\partial E(V)/\partial \theta) > 0$. Cela correspond à l'attitude de tout agent ayant des préférences pour la diversification quand il fait face à un risque bivarié négativement corrélé.

En résumé, d'autant moins élevé est π_{SI} ($\pi_{SI} \rightarrow 0$), d'autant plus élevée est la corrélation négative ($|\rho_{SI}| \rightarrow 1$) entre les risques d'inondation et de sécheresse, d'autant plus que l'exploitant est sûr de récupérer le fruit de son effort et d'autant plus important est l'effort d'arrosage consenti. La situation contraire s'observe quand $\pi_{SI} \rightarrow pq$. Ainsi, si $\pi_{SI} \in]0, pq[$, il existe une corrélation négative non nulle entre S et I . Par suite, l'exploitant consentira un niveau d'effort optimal $\theta^* \in]0, 1[$ tel que $(\partial E(V)/\partial \theta^*) = 0$. Ce cas intermédiaire va être analysé dans la sous-section suivante.

Impact formel de π_{SI} sur l'effort optimal d'arrosage. Les analyses précédentes suggèrent que d'autant plus élevé est π_{SI} ($\pi_{SI} \rightarrow pq$), d'autant plus élevée est la corrélation positive ($\rho_{SI} \rightarrow 1$) entre les risques d'inondation et de sécheresse, d'autant moins que l'exploitant est sûr de récupérer le fruit de son effort et d'autant moins important est l'effort d'arrosage consenti. On se propose ici d'évaluer formellement l'impact de π_{SI} et donc de la corrélation entre les risques de sécheresse et d'inondation sur l'effort d'arrosage. Soit θ^* le niveau optimal de θ déterminé à partir de la *CPO* obtenue à l'équation (2.5). En notant :

$$D_1 = (\alpha\Pi - c)^2 (p - \pi_{SI}) v'' (R + [\alpha\theta^* + (1 - \alpha)]\Pi - c\theta^*)$$

$$D_2 = c^2 \pi_{SI} v'' (R - c\theta^*),$$

l'impact de π_{SI} sur θ^* est tel que (Annexe 2.A.1.1, équation 2.52) :

$$\frac{d\theta^*}{d\pi_{SI}} = \frac{(\alpha\Pi - c) v' (R + (\alpha\theta^* + (1 - \alpha))\Pi - c\theta^*) + cv' (R - c\theta^*)}{D_1 + D_2} < 0$$

car $v'(\cdot) > 0$ et $v''(\cdot) < 0$.

Comme soulignée précédemment, l'aversion aux risques de sécheresse et d'inondation corrélés positivement décourage l'exploitant à investir de moins en moins d'effort d'arrosage à mesure que la probabilité qu'un épisode de sécheresse suivi de la survenue d'une inondation augmente.

2.3.1.1 Risque Relatif, effort d'arrosage et sévérité de la sécheresse

Comme on vient de le voir, l'effort d'arrosage est d'autant moins important que $\pi_{SI} \rightarrow p$, p étant la probabilité qu'il se produise une situation sécheresse. *La probabilité d'avoir une inondation sur la probabilité de ne pas en avoir une sachant qu'il y a sécheresse* - le risque relatif, noté ϱ par la suite - constitue un indicateur intéressant qui permet d'évaluer, dans une autre perspective, le comportement de l'exploitant en termes d'effort d'arrosage déployé. Dans le cas où $\pi_{SI} \in]0, pq[$, le risque relatif, donné par $\varrho(I/S) \equiv P(I/S)/P(I^c/S) = \pi_{SI}/(p - \pi_{SI})$, définit implicitement le niveau d'effort optimal $\theta^* \in$

$]0, 1[$ tel que $(\partial E(V)/\partial \theta) = 0$ (équation 2.5) par l'équation :

$$\varrho \equiv \frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{\alpha \Pi - c}{c} \right) \left(\frac{v'(R + [\alpha \theta + (1 - \alpha)] \Pi - c\theta)}{v'(R - c\theta)} \right). \quad (2.9)$$

à condition que $\alpha \Pi > c$. On peut remarquer qu'effectivement quand $\pi_{SI} \rightarrow pq$,

$$\varrho \equiv \frac{\pi_{SI}}{p - \pi_{SI}} \rightarrow \frac{q}{1 - q}.$$

La corrélation négative entre les risques d'inondation et de sécheresse est d'autant moins élevée quand $\pi_{SI} \rightarrow pq$. Par suite, d'autant plus élevée est la probabilité q d'avoir une inondation, d'autant plus élevé est ϱ , d'autant moins que l'exploitant est sûr de récupérer le fruit de son effort et donc d'autant moins important est l'effort d'arrosage consenti. Pour des niveaux de corrélation négative ρ_{SI} tels que $|\rho_{SI}| \rightarrow 0$ entre les risques d'inondation et de sécheresse, un agriculteur averse au risque fera un effort d'arrosage plus important et choisira de s'assurer contre le risque auquel il est capable de s'assurer (Finkelshtain et Scarsini, 1999). Pour R et Π donnés, et pour des valeurs du risque relatif ϱ , étudions maintenant le comportement du niveau d'effort optimum d'arrosage θ en fonction du degré de la sévérité de la sécheresse α .

Quand il n'y a pas de sécheresse. Quand la sécheresse est sans impact sur les profits de la récolte ($\alpha = 0$), l'exploitant n'est en rien incité à consentir un certain effort et pour cause, $\forall \theta \in [0, 1]$, l'équation (2.5) donne

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\alpha=0} = -\pi_{SI} c v'(R - c\theta) - c(p - \pi_{SI}) v'(R + \Pi - c\theta) < 0.$$

et donc, il n'a évidemment aucun intérêt à faire un effort d'arrosage, indépendamment du niveau de son revenu et de son attitude face au risque ($v'(\cdot) > 0$).

Quand la sécheresse est d'intensité maximale. Quand la sécheresse empêche toute récolte ($\alpha = 1$), le niveau d'effort optimal est tel que

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\alpha=1} = -\pi_{SI} c v'(R - c\theta) + (p - \pi_{SI}) [\Pi - c] v'(R + \theta\Pi - c\theta). \quad (2.10)$$

Dans ce cas, l'impact de l'effort déployé sur le niveau d'utilité de l'exploitant peut être différent dépendamment de si cet effort est maximal ou nul. Quand le niveau d'effort est maximal ($\theta = 1$), cette dernière expression s'écrit :

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\alpha=1}^{\theta=1} = -\pi_{SI} c v'(R - c) + (p - \pi_{SI}) [\Pi - c] v'(R + \Pi - c).$$

Et donc, pour que l'exploitant aie effectivement le désir d'effectuer cet effort maximal, il faut que

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\alpha=1}^{\theta=1} > 0$$

et par suite

$$\frac{\pi_{SI}}{p - \pi_{SI}} < \left(\frac{\Pi - c}{c} \right) \left(\frac{v'(R + \Pi - c)}{v'(R - c)} \right). \quad (2.11)$$

On retrouve le fait que, pour que l'agent ait envie de faire un effort d'arrosage, il faut que la probabilité avec laquelle son effort est vain (parce qu'il y a une inondation qui survient) soit relativement faible. En particulier si $\pi_{SI}/(p - \pi_{SI})$ est trop grand ($q \rightarrow 1$), il n'exercera pas un niveau d'effort maximal, même dans le cas des sécheresses les plus sévères. Si l'exploitant juge opportun de faire de l'effort, celui-ci sera toujours maximum ($\theta = 1$) quand la sévérité de la sécheresse est maximale : $\alpha = 1$.

Au cas où la sécheresse empêche toute récolte et que l'exploitant ne consent aucun effort ($\theta = 0$), l'incitation à arroser en cas de sécheresse maximale est donnée par :

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\alpha=1}^{\theta=0} = [-\pi_{SI} c + (p - \pi_{SI}) (\Pi - c)] v'(R) > 0.$$

Ce qui implique que :

$$\frac{\pi_{SI}}{p - \pi_{SI}} < \frac{\Pi - c}{c}. \quad (2.12)$$

Dans le cas contraire, aucun effort ne se fera fait c'est-à-dire si $[\partial E(V)/\partial \theta]_{\alpha=1}^{\theta=0} < 0$.

Par conséquent, quand la sécheresse est maximale, il y a deux seuils importants qui se dégagent pour le ratio $\pi_{SI}/(p - \pi_{SI})$. Si $[\pi_{SI}/(p - \pi_{SI})] > (\Pi - c)/c$, il n'y a jamais d'arrosage. Par contre, en regardant les inéquations (2.11) et (2.12) et sachant que $[v'(R + \Pi - c)/v'(R - c)] < 1$, il s'ensuit que

$$\left(\frac{v'(R + \Pi - c)}{v'(R - c)}\right) \left(\frac{\Pi - c}{c}\right) < \frac{\pi_{SI}}{p - \pi_{SI}} < \frac{\Pi - c}{c}. \quad (2.13)$$

Dans ce dernier cas, un effort d'arrosage sera fait mais il ne sera jamais maximum. Plus précisément, ce niveau d'effort est défini par, en résolvant la *CPO* (2.10)

$$\frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{v'(R + \theta\Pi - c\theta)}{v'(R - c\theta)}\right) \left(\frac{\Pi - c}{c}\right).$$

Niveaux de sécheresse « intermédiaires » ($0 < \alpha < 1$). Pour des niveaux de sécheresse intermédiaires, il est intéressant de caractériser la valeur maximale du risque relatif $\pi_{SI}/(p - \pi_{SI})$ en deçà de laquelle on veut toujours faire un effort d'arrosage maximum. Cette valeur du risque relatif est définie par

$$\left.\frac{\partial E(V)}{\partial \theta}\right|_{\theta=1} = -\pi_{SI}c v'(R - c) + (p - \pi_{SI})[\alpha\Pi - c] v'(R + \Pi - c) = 0.$$

On obtient :

$$\frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{v'(R + \Pi - c)}{v'(R - c)}\right) \left(\frac{\alpha\Pi - c}{c}\right). \quad (2.14)$$

De façon assez naturelle, à mesure que la sécheresse devient de plus en plus en plus sévère ($\alpha \rightarrow 1$), la probabilité d'avoir une inondation sur la probabilité de ne pas en avoir une sachant qu'il y a sécheresse- le risque relatif $\pi_{SI}/(p - \pi_{SI})$ - augmente. On peut déduire aussi de cette même équation que quand $\alpha \leq c/\Pi$, alors, on ne veut jamais faire l'effort maximal.

De la même manière, on peut caractériser la valeur maximale du risque relatif $\pi_{SI}/(p - \pi_{SI})$ en deçà de laquelle on ne veut pas faire d'effort d'arrosage ($\theta = 0$). Cette valeur est

donnée par

$$\left. \frac{\partial E(V)}{\partial \theta} \right|_{\theta=0} = -\pi_{SI} c v'(R) + (p - \pi_{SI}) [\alpha \Pi - c] v'(R + (1 - \alpha) \Pi) = 0.$$

On obtient cette fois :

$$\frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{v'(R + (1 - \alpha) \Pi)}{v'(R)} \right) \left(\frac{\alpha \Pi - c}{c} \right). \quad (2.15)$$

A nouveau, puisque $v'(R + (1 - \alpha) \Pi)$ est croissant en α ($v'(\cdot) > 0$), le cote caractérisé par (2.15) est donc croissant en α . Et bien sûr, quand $\alpha \leq c/\Pi$, c'est à dire quand les pertes dues à la sécheresse sont inférieures au coût de l'arrosage, on ne va jamais l'entreprendre.

Niveaux de sécheresse et d'effort d'arrosage intermédiaires ($0 < \alpha, \theta < 1$). De manière générale, pour $0 < \alpha, \theta < 1$, pourvu que la corrélation entre risque de sécheresse et risque d'inondation ne soit pas nulle ou positive, un effort d'arrosage sera fait mais il ne sera jamais maximal. Dans ce cas, le risque relatif ϱ est tel que :

$$\left(\frac{v'(R + \Pi - c)}{v'(R - c)} \right) \left(\frac{\alpha \Pi - c}{c} \right) < \frac{\pi_{SI}}{p - \pi_{SI}} < \left(\frac{v'(R + (1 - \alpha) \Pi)}{v'(R)} \right) \left(\frac{\alpha \Pi - c}{c} \right) \quad (2.16)$$

2.3.1.2 Impact des revenus sur les efforts d'arrosage

Pour analyser le lien entre l'effort d'arrosage et le niveau de revenu R , on suppose que l'aversion au risque décroît avec le revenu étant donnée la concavité de $v(\cdot)$. Cette hypothèse, analysée par Pratt (1964) implique que $a_{RA}(R)$ - le *coefficient d'aversion au risque absolu* - est une fonction décroissante du revenu si et seulement, pour chaque risque, l'équivalent certain d'un investisseur averse au risque est d'autant plus important que ses actifs le sont et que sa prime de risque et ce qu'il serait disposé à payer pour s'assurer contre le risque soient d'autant plus faibles (Jehle et Reny, 2011). Cette hypothèse est supportée empiriquement par un nombre de travaux de recherche tels, entre autres, les travaux de Bar-Shira et al. (1997), Riley Jr et Chow (1992) et Szpiro (1983).

L'impact du revenu sur l'effort d'arrosage est analysé en utilisant le *coefficient d'aver-*

sion au risque absolu introduit par Pratt (1964). Notons $a_{RA}(R)$ ce coefficient. Il est donné par :

$$a_{RA}(R) = \frac{-v''(R)}{v'(R)}. \quad (2.17)$$

Le niveau d'effort optimal, établi dans l'équation (2.5), est caractérisé par l'équation implicite (2.9). Evaluons l'impact du niveau de revenu R sur l'effort d'arrosage. En dérivant la *CPO* de l'équation (2.5) par rapport à R , on établit que (Annexe 2.A.1.2) :

$$\frac{d\theta}{dR} = \frac{a_{RA}(R - c\theta) - a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta) + (\alpha\Pi - c)a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}. \quad (2.18)$$

En adoptant l'hypothèse de décroissance de l'aversion au risque avec le revenu, on a alors

$$a_{RA}(R - c\theta) \geq a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)$$

et donc $d\theta/dR \geq 0$ pourvu que $\alpha\Pi > c$. En d'autres termes, en situation de sécheresse, l'exploitant est davantage incité à arroser quand son revenu augmente sachant que la corrélation entre risques de sécheresse et d'inondation est négative.

Les circonstances où l'on ne fait aucun effort d'arrosage (telles que caractérisées par l'équation 2.15) diminuent quand le revenu augmente. En effet, comme (Annexe 2.A.1.2)

$$\frac{d\varrho}{dR} \equiv \frac{d}{dR} \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) = [a_{RA}(R) - a_{RA}(R + (1 - \alpha)\Pi)] \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) \geq 0$$

donc le niveau du risque relatif (équation 2.15) à partir duquel on ne fait aucun effort d'arrosage augmente, augmentation induite probablement par une corrélation positive significativement élevée entre risques de sécheresse et d'inondation.

De même, les circonstances où l'on fait l'effort maximum (telles que caractérisées par l'équation 2.14) sont respectivement plus ou moins fréquentes quand le revenu augmente. En effet, comme (Annexe 2.A.1.2)

$$\frac{d\varrho}{dR} \equiv \frac{d}{dR} \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) = [a_{RA}(R - c) - a_{RA}(R + \Pi - c)] \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) \geq 0$$

et donc le niveau du risque relatif - en combinant avec l'équation (2.14) - en deçà duquel on fait toujours l'effort maximal d'arrosage augmente.

Sous l'hypothèse que l'aversion au risque décroît avec le revenu, une amélioration des revenus de l'exploitant constitue donc un élément qui l'incite à investir davantage d'effort d'arrosage dans une situation de sécheresse telle que les risques de sécheresse et d'inondation sont non positivement corrélés.

2.3.2 Revenu espéré avec et sans arrosage

Le petit exploitant agricole peut continuer à cultiver sa parcelle en s'abandonnant aux conditions climatiques ou tout au contraire décider de s'assurer contre le risque de sécheresse en arrosant. Si l'on tient compte du coût de l'arrosage et du degré d'aversion au risque de l'agriculteur, quel serait pour l'exploitant, le choix le plus rentable ?

De manière générale, l'arrosage contribue à faire augmenter la productivité agricole, le revenu et la consommation des ménages. Différentes études réalisées confirment une relation positive entre arrosage à petite échelle et amélioration des revenus et la productivité agricole (Mango et al., 2018). Ces conclusions sont aussi confirmées par Dillon (2011), l'impact sur la consommation percapita étant plus importante dans le cas des systèmes d'irrigation à grande échelle. Les potentialités de l'arrosage à petite échelle sont positivement évaluées à travers une recherche publiée par Xie et al. (2014) mettant en avant une amélioration des revenus dépendamment de la technologie d'arrosage utilisée.

En notant θ^* le niveau d'effort d'arrosage choisi par l'exploitant d'après la règle donnée à l'équation (2.5) quand il fait de l'arrosage, son revenu espéré $E(R_a)$, étant donné l'ensemble des états de la nature, est donné par :

$$\begin{aligned}
E(R_a) &= \pi_{SI}(R - c\theta^*) + (q - \pi_{SI})R \\
&+ (p - \pi_{SI})[R + (\alpha\theta^* + (1 - \alpha)\Pi) - c\theta^*] \\
&+ (1 - p - q + \pi_{SI})(R + \Pi) \\
&= R + (1 - q)\Pi + (p - \pi_{SI})(\theta^* - 1)\alpha\Pi - pc\theta^*.
\end{aligned} \tag{2.19}$$

Quand $0 < \alpha < 1$, l'effort d'arrosage est maximal ($\theta^* = 1$) si (condition établie à l'équation 2.11)

$$\frac{\pi_{SI}}{p - \pi_{SI}} \leq \left(\frac{v'(R + \Pi - c)}{v'(R - c)} \right) \left(\frac{\alpha\Pi - c}{c} \right).$$

Le revenu espéré est alors donné par

$$E(R_a) = R + (1 - q)\Pi - pc. \quad (2.20)$$

Naturellement, le revenu espéré de l'exploitant diminue d'autant plus que la probabilité qu'il survienne une inondation et/ou une sécheresse augmentent. Un épisode de sécheresse fait perdre les coûts liés à l'effort d'arrosage de l'ordre de pc et la survenue d'une inondation impactera négativement le profit de l'exploitant de l'ordre de $q\Pi$.

Si au contraire

$$\frac{\pi_{SI}}{p - \pi_{SI}} \geq \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \left(\frac{\alpha\Pi - c}{c} \right),$$

(condition définie à partir de la *CPO* donnant l'équation 2.15), alors il n'y a pas d'arrosage.

On a alors ($\theta^* = 0$) et donc :

$$E(R_{\bar{a}}) = R + [(1 - q) - \alpha(p - \pi_{SI})]\Pi. \quad (2.21)$$

Le revenu espéré de l'agriculteur diminuera de $q\Pi$ s'il se produira une inondation et de $\alpha p\Pi$ si une sécheresse d'intensité α survient. En effet, quand l'exploitant n'arrose pas, avec la survenue d'une sécheresse, d'autant plus intense est cette sécheresse ($\alpha \rightarrow 1$), d'autant plus importante est la diminution de son revenu espéré. La diminution du revenu espéré est d'autant plus importante que la corrélation entre les risques de sécheresse et d'inondation est positive ($\pi_{SI} \rightarrow p$) ou négligeable ($\pi_{SI} \rightarrow pq$). Cette situation résulte de l'aversion aux risques positivement corrélés de sécheresse et d'inondation de l'agriculteur le décourageant d'investir dans l'activité d'arrosage.

L'exploitant bénéficiera d'un différentiel positif de revenu espéré ($E_d(R)$) quand il fait

de l'arrosage en cas de sécheresse. Ce différentiel de revenu espéré donné par

$$E_d(R) \equiv E(R_a) - E(R_{\bar{a}}) = \alpha(p - \pi_{SI})\Pi - pc \quad (2.22)$$

est avantageux à condition que

$$\begin{aligned} \alpha(p - \pi_{SI})\Pi - pc &> 0 \\ \frac{\pi_{SI}}{p} &< 1 - \frac{c}{\alpha\Pi} \end{aligned} \quad (2.23)$$

c'est à dire, quand il y a sécheresse, la probabilité qu'il y ait une d'inondation qui succède à un épisode de sécheresse n'atteigne pas la fraction de profit réalisée, pour avoir déployé un certain effort proportionnel à l'intensité de la sécheresse.

Dans les cas intermédiaires, c'est à dire quand

$$\left(\frac{v'(R + \Pi - c)}{v'(R - c)}\right) \left(\frac{\alpha\Pi - c}{c}\right) < \frac{\pi_{SI}}{p - \pi_{SI}} < \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)}\right) \left(\frac{\alpha\Pi - c}{c}\right), \quad (2.24)$$

le niveau d'arrosage optimum est défini par l'équation (2.9) et le revenu espéré par (2.19).

2.3.3 Effort d'arrosage en présence d'assurance sur les inondations.

Dans le cas où un marché d'assurance existe, les agents économiques averses au risque achètent une assurance contre le paiement d'une prime pour se prémunir contre des pertes financières potentielles mais incertaines. Ce contrat d'assurance offre à l'assuré la possibilité de réduire ses pertes financières en mutualisant le risque à travers l'entreprise d'assurance. Ce mécanisme s'avère très difficile d'application dans le cas des risques liés aux changements climatiques (CC) d'intensité de plus en plus importante et entretenant de plus en plus d'incertitudes. Le marché des capitaux, à son niveau de développement actuel n'arrive pas à mettre à la disposition des personnes qui le souhaitent des produits d'assurance contre les risques climatiques (Doherty, 1997). A coté des problèmes classiques de sélection adverse et d'hasard moral, les potentiels assurés n'ont pas le contrôle de l'événement

nement créant le risque, le plus souvent de grande ampleur et dont l'importance et les difficultés d'estimation des probabilités associées empêchent le développement des marchés d'assurance appropriés (Jaffee et Russell, 1997). Ces difficultés proviennent de l'ampleur des risques liés aux *CC*, dans certains cas de l'inexistence de marché d'assurance pour offrir de tels produits, la corrélation géographique rendant impossible la diversification via le théorème de la limite centrale, etc.(Charpentier, 2008). Dans ce contexte de *CC*, l'inondation reste l'un des phénomènes les plus fréquents observés durant ces dernières années. La problématique d'une assurance-inondation, avec un taux de pénétration de 40% dans les pays développés et 5% dans les PED, reste posée dans les PED (Surminski, 2013) en raison de multiples facteurs tels que présentés par Charpentier (2008). L'importance et la croissance des risques liées aux rapides changements climatiques et les pertes qui en résultent soulèvent et entretiennent la nécessité d'élaborer des outils adéquats de transferts de risques liés à ces *CC*, en particulier le risque d'inondation, surtout dans les pays à faible revenu (PFR).

Dans le cas d'un petit exploitant, en dépit des efforts qu'il pourrait déployer pour arroser sa parcelle en cas de sécheresse, la corrélation existant entre le risque d'inondation (un risque covariant) et le risque de sécheresse milite pour la mise sur pied d'un outil permettant de faire face à cette situation. On se propose dans le cadre de ce modèle d'évaluer dans un premier temps l'impact d'une assurance exogène (une aide financière) en situation d'inondation et une couverture d'assurance-inondation endogène sur l'effort d'arrosage.

2.3.3.1 Impact d'une assurance exogène sur les inondations

Différentes recherches réalisées justifient une intervention de l'état pour venir en aide aux exploitants agricoles en cas de catastrophes naturels. Les imperfections du marché des capitaux rendant difficile la mise en place de produits d'assurance contre les risques climatiques (Miranda et Glauber, 1997; Skees et Reed, 1986; Chambers, 1989) constituent les principaux arguments en faveur d'une intervention de l'état. Théoriquement, des produits d'assurance contre les risques climatiques systémiques et covariants peuvent

être conçus mais les problèmes aigus de non disponibilité d'information complète et de manque d'incitations adéquates ne permettent pas d'avoir des produits pareto-optimaux (Nelson et Loehman, 1987). Des solutions de deuxième rang tels par exemple les subsides octroyés par l'état jouent un rôle important dans la mise en place et la réussite d'un système d'assurance contre les risques climatiques (Mahul et Stutley, 2010).

On suppose alors qu'en cas d'inondation, l'agriculteur reçoit une aide financière A en fonction de l'effort déployé.

Avec l'octroi potentiel d'une assistance financière A en cas d'inondation, l'utilité espérée de l'exploitant définie à l'équation (2.2), se ré-écrit :

$$\begin{aligned} E(V) &= \pi_{SI}v(R - c\theta + A) + (q - \pi_{SI})v(R + A) \\ &+ (p - \pi_{SI})v(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &+ (1 - p - q + \pi_{SI})v(R + \Pi). \end{aligned} \quad (2.25)$$

La condition du premier ordre par rapport à θ est telle que

$$\begin{aligned} 0 &= \frac{dE(V)}{d\theta} \equiv -\pi_{SI}cv'(R - c\theta + A) \\ &+ (p - \pi_{SI})[\alpha\Pi - c]v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta). \end{aligned} \quad (2.26)$$

En dérivant par rapport à A la CPO en θ donnée à l'équation précédente, on établit que (*Annexe 2.A.1.3*) :

$$\frac{d\theta}{dA} = \frac{\pi_{SI}cv''_A(\cdot)}{\pi_{SI}c^2v''_A(\cdot) + (p - \pi_{SI})(\alpha\Pi - c)^2v''_A(\cdot)} > 0. \quad (2.27)$$

où

$$\begin{aligned} v''_A(\cdot) &= v''(R - c\theta + A) \\ v''_A(\cdot) &= v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \end{aligned}$$

La positivité de $d\theta/dA$ confirme que telle qu'établie dans la littérature, la présence d'une aide financière exogène en cas d'inondation permettant à l'agriculteur de récupérer le fruit de son effort d'arrosage, l'incite à faire plus d'effort.

Supposons maintenant que la compensation financière A est conçue de sorte qu'elle couvre *l'intégralité des pertes* en cas d'inondation c'est-à-dire $A = \Pi$. Même en cas d'un effort maximal ($\theta = 1$), tous les risques liés à l'effort d'arrosage ne peuvent être pas être couverts. En effet, en faisant $A = \Pi$ dans l'équation (2.25), l'utilité espérée se ré-écrit :

$$\begin{aligned} E(V) &= \pi_{SI}v(R - c\theta + \Pi) + (q - \pi_{SI})v(R + \Pi) \\ &+ (p - \pi_{SI})v(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &+ (1 - p - q + \pi_{SI})v(R + \Pi). \end{aligned}$$

Dans le cas d'un effort d'arrosage maximal, l'utilité espérée quand la compensation couvre l'intégralité des pertes est donnée par :

$$E(V) = pv(R + \Pi - c) + (1 - p)v(R + \Pi) : \quad (2.28)$$

il y a donc un risque résiduel non assurable lié à la récupération des coûts d'arrosage c investis : *il est donc impossible d'avoir une couverture exogène complète*. Ce risque résiduel non assurable impacte négativement l'utilité espérée de l'exploitant même en cas d'un effort d'arrosage maximal :

$$\left. \frac{dE(V)}{dc} \right|_{A=\Pi, \theta=1} = -cpv'(R + \Pi - c) < 0.$$

Aide financière exogène et taxe de compensation en absence d'inondation.

Supposons à présent que l'aide financière A en cas d'inondation est octroyée au début de la période. Dans les situations où il n'y a pas d'inondation, le gouvernement impose sur le profit réalisé, une taxe $\tau_A \Pi \leq A$ telle que $0 < \tau_A < 1$ à des fins de compensation . On postule que l'effort de l'exploitant sera influencé par τ_A de sorte que $\theta = \theta(\tau_A)$. L'utilité

espérée du petit agriculteur devient :

$$\begin{aligned}
E(V) &= \pi_{SI}v(R - c\theta + A) + (q - \pi_{SI})v(R + A) \\
&+ (p - \pi_{SI})v\{R + [\alpha\theta + (1 - \alpha) - \tau_A]\Pi - c\theta\} \\
&+ (1 - p - q + \pi_{SI})v[R + (1 - \tau_A)\Pi].
\end{aligned} \tag{2.29}$$

Dans ce cas, l'effort maximal optimal θ qui maximise l'utilité espérée de l'exploitant est telle que :

$$\begin{aligned}
\frac{dE(V)}{d\theta} &= -\pi_{SI}cv'(R - c\theta + A) \\
&+ (p - \pi_{SI})(\alpha\Pi - c)v'\{R + [\alpha\theta + (1 - \alpha) - \tau_A]\Pi - c\theta\}.
\end{aligned} \tag{2.30}$$

En notant

$$\begin{aligned}
v''_{\frac{R}{c}}(\cdot) &= v''(R - c\theta + A), \\
v''_{\tau}(\cdot) &= v''[R + (\alpha\theta + (1 - \alpha) - \tau_A)\Pi - c\theta],
\end{aligned}$$

l'impact de cette taxe τ_A sur l'effort optimal d'arrosage est telle que (Annexe : équation 2.A.1.3)

$$\frac{d\theta}{d\tau_A} = \frac{(p - \pi_{SI})(\alpha\Pi - c)\Pi v''_{\tau}(\cdot)}{c^2\pi_{SI}cv''_{\frac{R}{c}}(\cdot) + (p - \pi_{SI})(\alpha\Pi - c)^2 v''_{\tau}(\cdot)} > 0. \tag{2.31}$$

A condition que $\Pi > c\theta$, l'impact de la taxe sur l'effort optimal est positif car $v''(\cdot) < 0$. Donc, même en imposant une taxe $\tau_A\Pi$ à des fins de compensation quand il n'y a pas d'inondation, l'exploitant est toujours incité à fournir, avec l'octroi d'une aide financière A , un effort d'arrosage en cas de sécheresse toutes les fois que la corrélation sécheresse-inondation est négative.

2.3.3.2 Couverture d'assurance endogène

Divers arguments sont avancés dans le cadre de la théorie de l'assurance pour expliquer la disponibilité partielle ou la non disponibilité de produits d'assurance pour couvrir les risques climatiques. De manière générale, la non disponibilité de l'information complète et l'inadéquation des incitations envers l'assuré et l'entreprise d'assurance empêchent l'obtention des équilibres pareto-optimaux sur le marché de l'assurance. De manière spécifique, les problèmes de sélection adverse - degré de risque de l'assuré non observable par l'assureur (asymétrie d'information) le portant à couvrir les assurés les plus risqués - et d'hasard moral - actions de l'assuré non observables par l'assureur et qui augmentent la probabilité que l'événement assuré se réalise - constituent les deux principaux problèmes en assurance (Nelson et Loehman, 1987; Chambers, 1989). Dans le cas des risques climatiques, le caractère covariant et non diversifiable de ces risques s'ajoutant aux problèmes traditionnels de sélection adverse et d'hasard moral, rend difficile la provision de services d'assurance pour de tels risques. Ce qui justifie l'intervention de l'Etat dans la provision de tels services et/ou la mise en place des incitations appropriées pour encourager la participation du secteur privé dans une telle entreprise (Mahul et Stutley, 2010).

Dans le cas des pays développés, différentes alternatives de couverture sont offertes aux producteurs agricoles. De ces alternatives, l'intervention du gouvernement sous différentes formes reste l'une des formes de protection les plus utilisées en raison du caractère covariant et non diversifiable des risques climatiques (Glauber et Peter, 2002; Nelson et Loehman, 1987). Les formes d'intervention vont de l'otroi direct d'une assistance financière aux victimes d'aléas climatiques, de la mise en place d'institutions offrant des couvertures d'assurance contre les risques climatiques, à des incitations appropriées au secteur privé d'assurance ou à l'établissement de partenariat public-privé pour la provision de services d'assurance contre de tels risques. Ces interventions, dans le cas où elles existeraient dans les pays en développement, sont très limitées et non systématiques. Cela est dû, entre autres, à la carence d'infrastructures d'assurance dans ces pays, à l'accès limité au marché de réassurance international, à l'absence d'une culture assurantielle et à une capacité limitée des producteurs agricoles d'évaluer l'ampleur des risques - et leur impact potentiel

- pesant sur leurs activités (Mahul et Stutley, 2010).

Dans le cas des pays sous-développés, l'agriculture occupe une place importante en termes de participation au produit intérieur brut et de l'emploi. Dans ces pays, les activités agricoles sont en majorité menées par des petits exploitants en général laissés à leur propre sort. Le taux de pauvreté est assez souvent très élevé au sein ce groupe d'exploitants agricoles. On se propose, dans le cadre de ce modèle, d'évaluer les *bénéfices et l'utilité* d'une police assurance inondation, ceteris paribus, pour un petit agriculteur qui, pour s'en sortir, décide d'investir un certain effort d'arrosage privé en cas de sécheresse.

On suppose alors que le petit agriculteur est assuré en cas d'inondation en achetant un seul contrat d'assurance. Sa couverture d'assurance est donnée par $C_{\rho A} = \rho A$ où $A \in [0, \Pi]$ représente son niveau de couverture, *dépendant de son niveau d'effort* θ et $0 < \rho < 1$ le taux de la prime d'assurance. Son utilité espérée s'écrit alors :

$$\begin{aligned} E(V) &= \pi_{SI} v(R - \rho A - c\theta + A) + (q - \pi_{SI}) v(R - \rho A + A) \\ &+ (p - \pi_{SI}) v(R - \rho A + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &+ (1 - p - q + \pi_{SI}) v(R - \rho A + \Pi). \end{aligned} \quad (2.32)$$

La *CPO* en θ déterminée à partir de l'équation (2.32) donne

$$\begin{aligned} \frac{\partial E(V)}{\partial \theta} &= -\pi_{SI} c v'(R - \rho A - c\theta + A) \\ &+ (p - \pi_{SI}) [\alpha\Pi - c] v'(R - \rho A + [\alpha\theta + (1 - \alpha)]\Pi - c\theta). \end{aligned} \quad (2.33)$$

La *CPO* en A , toujours à partir de (2.32), est donnée par

$$\begin{aligned} \frac{\partial E(V)}{\partial A} &= \pi_{SI} (1 - \rho) v'(R - \rho A - c\theta + A) \\ &+ (q - \pi_{SI}) (1 - \rho) v'(R - \rho A + A) \\ &- (p - \pi_{SI}) \rho v'(R - \rho A + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &- (1 - p - q + \pi_{SI}) \rho v'(R - \rho A + \Pi). \end{aligned} \quad (2.34)$$

Le niveau optimal d'effort à investir et de couverture d'assurance défini par la paire optimale (θ, A) , s'obtient par la solution du système :

$$\left\{ \begin{array}{l} (p - \pi_{SI}) [\alpha\Pi - c] v' (R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad = \pi_{SI} c v' (R - \rho A - c\theta + A) \quad (CPO \text{ en } \theta) \\ \pi_{SI} (1 - \rho) v' (R - \rho A - c\theta + A) + (q - \pi_{SI}) (1 - \rho) v' (R - \rho A + A) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad = (p - \pi_{SI}) \rho v' (R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\ \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad + (1 - p - q + \pi_{SI}) \rho v' (R - \rho A + \Pi) \quad (CPO \text{ en } A) \end{array} \right. \quad (2.35)$$

Niveau d'effort et couverture d'assurance. On a pu établir à l'équation (2.27) qu'une aide financière exogène A incite l'agriculteur à investir plus d'effort d'arrosage en situation de sécheresse et cela même quand une taxe sur le profit réalisé est imposée (équation 2.31). Evaluons maintenant comment se comporte l'exploitant en termes d'effort d'arrosage investi θ quand il contracte formellement une police d'assurance A pour se couvrir en cas d'inondation. Pour ce faire, on dérive l'équation (2.34) par rapport à θ . En adoptant les notations suivantes :

$$\begin{aligned} v''_{\theta I} &= v'' (R - \rho A + A) \\ v''_{SI} &= v'' (R - \rho A - c\theta + A) \\ v''_{S\theta} &= v'' (R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\ v''_{\theta\theta} &= v'' (R - \rho A + \Pi) \\ \Delta_{\theta A} &= (1 - \rho)^2 [\pi_{SI} v''_{SI} + (q - \pi_{SI}) v''_{\theta I}] \\ &\quad + \rho^2 [(p - \pi_{SI}) v''_{S\theta} + (1 - p - q + \pi_{SI}) v''_{\theta\theta}], \end{aligned}$$

on a pu établir que (Annexe 2.A.1.4) :

$$\frac{dA}{d\theta} = \frac{\pi_{SI}}{\Delta_{\theta A}} c (1 - \rho) v''_{SI} + \frac{p - \pi_{SI}}{\Delta_{\theta A}} (\alpha\Pi - c) \rho v''_{S\theta}. \quad (2.36)$$

Etant donné que $v''(\cdot) < 0$, on n'en déduit que $\Delta_{\theta A} < 0$ et donc $(dA/d\theta) > 0$. En effet, une assurance inondation, en couvrant le risque d'inondation et les pertes potentielles de revenu associées, met donc en confiance l'agriculteur ayant une forte aversion au risque (Dionne et Eeckhoudt, 1985; Briys et Schlesinger, 1990) et averse aux risques positivement corrélés pour investir davantage d'effort en arrosant sa parcelle.

Aucune couverture d'assurance. Quand l'agriculteur n'est pas sûr de récupérer le fruit de son effort d'arrosage, il n'arrose pas. Dans ce cas de figure, il n'est ni intéressant ni rentable de contracter une couverture d'assurance. Dans le cas où il consent un effort d'arrosage, l'exploitant averse au risque est incité à contracter une couverture d'assurance inondation tant que la prime d'assurance ne dépasse pas sa prime de risque - le montant d'argent qu'il est disposé à payer pour se prémunir contre le risque d'inondation. Dans le cas contraire, il n'achètera pas d'assurance et donc $(dE(V)/dA) < 0$. Autrement dit, quand la couverture d'assurance inondation impacte négativement l'utilité de l'exploitant, il n'est plus intéressé à s'assurer et donc $A = 0$. Ainsi, le coût unitaire ρ de la couverture d'assurance, déterminé à partir de la CPO (2.34), est tel que $(dE(V)/dA) < 0$. Ce qui implique (Annexe 2.A.1.4)

$$\rho_{A=0} > \frac{\pi_{SI}}{\Delta'_\rho} v'(R - c\theta) + \frac{q - \pi_{SI}}{\Delta'_\rho} v'(R) \equiv \rho_m \quad (2.37)$$

avec

$$\begin{aligned} \Delta'_\rho &= (p - \pi_{SI}) v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &\quad + (1 - p - q + \pi_{SI}) v'(R + \Pi) \\ &\quad + \pi_{SI} v'(R - c\theta) + (q - \pi_{SI}) v'(R) \end{aligned}$$

Tant que $\rho > \rho_m$, le coût unitaire étant trop élevé, ne permet pas de compenser la perte d'utilité qui résulte de l'achat d'une telle police d'assurance : aucune couverture d'assurance n'est contractée.

Couverture d'assurance actuariellement juste. La couverture d'assurance A est rentable si la prime (coût de l'assurance) ρA vaut au moins l'indemnité espérée par l'assuré $E(A) = qA$, A est le montant de l'indemnité versé à l'assuré en cas d'inondation. La couverture d'assurance A est dite alors actuariellement juste ou non chargée quand la prime unitaire d'assurance est tout juste égale à l'espérance de l'indemnité soit $\rho = q$.

Assurance endogène, effort d'arrosage maximal et prime actuarielle. On considère le cas où l'agriculteur fournit un niveau d'effort d'arrosage maximal ($\theta = 1$) et que la prime d'assurance unitaire est actuarielle ($\rho = q$). De ce fait, la *CPO* en A à l'équation (2.34) se ré-écrit :

$$0 = \frac{\partial E(V)}{\partial A} = (1 - q) [\pi_{SI} v'(R - qA + A - c) + (q - \pi_{SI}) v'(R - qA + A)] - q [(p - \pi_{SI}) v'(R - qA + \Pi - c) + (1 - p - q + \pi_{SI}) v'(R - qA + \Pi)]. \quad (2.38)$$

Supposons que l'exploitant contracte une couverture d'assurance A équivalente à l'intégralité des profits Π réalisés. L'hypothèse $A = \Pi$ s'interprète comme la contraction d'une couverture $C_{\rho A}$ complète par l'agriculteur. Alors la couverture optimale est telle que

$$0 = \left. \frac{\partial E(V)}{\partial A} \right|_{A=\Pi} = [(1 - q) \pi_{SI} - q(p - \pi_{SI})] v'(R - qA + A - c) - [q(1 - p - q + \pi_{SI}) - (1 - q)(q - \pi_{SI})] v'(R - qA + A).$$

Ce qui donne

$$v'(R - qA + A - c) = v'(R - qA + A). \quad (2.39)$$

Donc, la couverture $C_{\rho A}$ serait complète que si $c = 0$. En situation d'un effort maximal investi, le coût d'arrosage associé ne peut être nul : il n'est donc pas possible d'avoir une couverture complète du fait de l'existence d'un risque résiduel non assurable lié au coût unitaire c de l'effort d'arrosage. On peut voir qu'à partir de la *CPO* donnée en (2.38),

l'impact du coût unitaire d'arrosage c est tel que

$$\begin{aligned} \frac{d}{dc} \left[\frac{\partial E(V)}{\partial A} \right] &= -(1-q) \pi_{SI} v'' (R - qA + A - c) \\ &\quad + q(p - \pi_{SI}) v'' (R - qA + \Pi - c). \end{aligned}$$

Sous l'hypothèse que $A = \Pi$,

$$\frac{d}{dc} \left[\frac{\partial E(V)}{\partial A} \right] = (pq - \pi_{SI}) v'' [R + (1-q)\Pi - c]. \quad (2.40)$$

Par suite, si $pq > \pi_{SI}$, alors

$$\frac{d}{dc} \left[\frac{\partial E(V)}{\partial A} \right] < 0 :$$

le risque résiduel induit par le coût d'arrosage tend à réduire la couverture d'assurance pour l'exploitant averse aux risques de sécheresse et d'inondation positivement corrélés ($pq > \pi_{SI} \implies \rho_{SI} > 0$). Par contre si $\pi_{SI} \rightarrow \min\{p, q\}$ donc $pq < \pi_{SI}$ ($\implies \rho_{SI} < 0$) alors

$$\frac{d}{dc} \left[\frac{\partial E(V)}{\partial A} \right] > 0 :$$

la couverture d'assurance augmente avec le coût d'arrosage.

Supposons maintenant que l'agriculteur décide de contracter une assurance A couvrant l'intégralité de ses profits et les coûts c investis pour un effort maximal c'est-à-dire $A = \Pi + c$. La *CPO* (2.38) devient (Annexe 2.A.1.4) ,

$$\begin{aligned} \frac{\partial E(V)}{\partial A} &= q(p - \pi_{SI}) [v'(R + (1-q)\Pi - qc) \\ &\quad - v'(R + (1-q)\Pi - (1+q)c)] \\ &\quad + (1-q)(q - \pi_{SI}) [v'(R + (1-q)(\Pi + c)) \\ &\quad - v'(R + (1-q)\Pi - qc)] \\ &> 0. \end{aligned} \quad (2.41)$$

En conséquence, une couverture de $A = \Pi + c$ contribue encore à faire augmenter l'utilité

de l'exploitant : d'autres risques pesant sur le coût d'arrosage investi incite l'agriculteur à vouloir couvrir non seulement ses profits mais aussi les coûts d'arrosage encourus. Cette attitude relève d'un comportement prudent de l'exploitant qui a tendance à se surprotéger en présence de risques inhérents.

Proposition 2.1. *Il n'est pas possible de trouver une paire (θ, A) qui permet d'éliminer tout risque.*

Démonstration. Voir (2.39) , (2.40) et (2.41). □

Pas d'arrosage et pas d'assurance. Quand, dans une situation de sécheresse d'intensité quelconque α ($0 < \alpha < 1$), il n'y a ni effort d'arrosage ($\theta = 0$: CPO 2.33) ni d'assurance ($A = 0$: CPO 2.34), l'agriculteur fait face à un ensemble de situations sans espoir. Les situations où il est *impossible* d'avoir une assurance inondation rentable et où l'exploitant *ne fait jamais l'effort d'arroser* sont donc telles que

$$\left\{ \begin{array}{l} \left. \frac{dE(V)}{d\theta} \right|_{A=\theta=0} < 0 \\ \left. \frac{dE(V)}{dA} \right|_{A=\theta=0} < 0 \end{array} \right. .$$

Ces deux conditions sont vérifiées (*Annexe : 2.A.1.4*) quand :

$$\begin{aligned} \left(\frac{1-q}{p} \right) \left[1 + \left(\frac{v'(R + (1-\alpha)\Pi)}{v'(R)} \right) \left(\frac{\alpha\Pi - c}{c} \right) \right] &< \frac{1-q}{p - \pi_{SI}} \\ &< \frac{[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} < 1 \end{aligned} \quad (2.42)$$

En particulier, ces deux conditions sont valides si $p + q > 1$. Par conséquent, pour qu'il soit possible d'avoir une assurance inondation rentable et où l'exploitant fait un certain effort d'arrosage, il faut que la probabilité d'avoir une sécheresse soit au plus égale à la probabilité de ne pas avoir une inondation ($p + q \leq 1$).

Proposition 2.2. *La condition 2.42 est vérifiée tant que $q \in]1 - p, 1]$.*

Démonstration. Annexe 2.A.1.5

Les situations où il est *impossible* d'avoir une assurance inondation rentable et où l'exploitant *ne fait jamais l'effort d'arroser*, sont sans espoir. Ces situations correspondent

naturellement à celles où il est de plus en plus probable qu'il survienne une inondation, quand il apparaît de moins en moins probable qu'il s'installe une sécheresse, décourageant ainsi le moindre effort d'arrosage. \square

2.3.3.3 Détermination simultanée d'une couverture d'assurance A et d'un niveau d'effort d'arrosage θ optimaux

La couverture d'assurance optimale A et le niveau d'effort optimum θ sont conjointement déterminés par la solution intérieure pour θ et A , si elle existe, du système formé par les équations (2.33) et (2.34). De la CPO (2.33) par rapport à θ , on en déduit :

$$\frac{c}{\alpha\Pi - c} \pi_{SI} v'(R - \rho A - c\theta + A) = (p - \pi_{SI}) v'(R - \rho A + [\alpha\theta + (1 - \alpha)]\Pi - c\theta).$$

Et, en remplaçant cette dernière expression dans la CPO (2.34), on obtient

$$\begin{aligned} \frac{\partial E(V)}{\partial A} &= (q - \pi_{SI}) (1 - \rho) v'(R - \rho A + A) \\ &+ \left[1 - \rho \left(\frac{\alpha\Pi}{\alpha\Pi - c} \right) \right] \pi_{SI} v'(R - \rho A - c\theta + A) \\ &- (1 - p - q + \pi_{SI}) \rho v'(R - \rho A + \Pi). \end{aligned} \quad (2.43)$$

La solution de l'équation précédente donne le niveau d'assurance optimal A^* garantissant la rentabilité de l'assurance pour un niveau d'effort optimal θ^* dans l'ensemble des situations décrites en (2.1).

Proposition 2.3. *Sous les hypothèses du modèle, il existe une solution intérieure (θ^*, A^*) qui maximise l'utilité espérée de l'exploitant agricole.*

Démonstration. Annexe 2.A.2.2. \square

2.4 Illustrations des principaux résultats

On adopte une fonction d'utilité de type *CRRA* (*Constant Relative Risk Aversion*) pour illustrer les résultats en raison de son importante utilisation dans la littérature, sa simplicité et ses propriétés intéressantes (Wakker, 2008; Holt et Laury, 2002; Palacios-Huerta et Serrano, 2006). En particulier, cette fonction d'utilité satisfait la condition de Inada (1963) stipulant qu'il est toujours plus utile de consommer au moins une unité d'un bien essentiel que de ne pas en consommer.

Une fonction d'utilité de type *CRRA* est définie par :

$$u(w) = \begin{cases} \frac{w^{1-\eta}}{1-\eta}, & w > 0 \text{ et } 1 \neq \eta > 0, \\ \ln w, & w > 0 \text{ et } \eta = 1. \end{cases} \quad (2.44)$$

où w est le niveau de richesse et η le coefficient d'Aversion au Risque Relatif (constant) d'Arrow-Pratt (Pratt, 1964) tel que $\forall w, \eta \equiv a_{RR}(w) = -w [u''(w) / u'(w)]$. Le coefficient d'Aversion au Risque Absolu, a_{RA} , est défini par : $a_{RA}(w) = \eta/w, w > 0$.

2.4.1 Sécheresse d'intensité maximale, sécheresse/inondation et effort d'arrosage

Quand la sécheresse est d'intensité maximale ($\alpha = 1$), l'expression générale du risque relatif ϱ définie à partir de l'équation (2.9) pour $0 < \theta < 1$ s'écrit :

$$\varrho \equiv \frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{R + \theta\Pi - c\theta}{R - c\theta} \right)^{-\eta} \left(\frac{\Pi - c}{c} \right) \quad (2.45)$$

tel que $R > c\theta$ et $\alpha\Pi > c$. Ce qui permet d'avoir explicitement

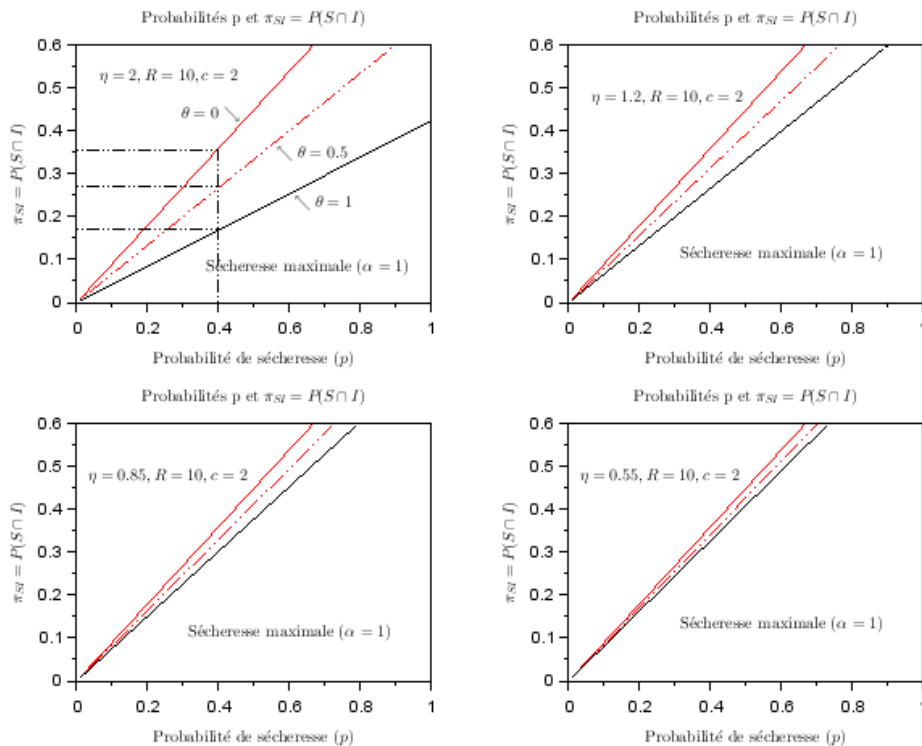
$$\pi_{SI} = \frac{p \left[\left(\frac{\Pi - c}{c} \right) \left(\frac{R + \theta\Pi - c\theta}{R - c\theta} \right)^{-\eta} \right]}{1 + \left(\frac{\Pi - c}{c} \right) \left(\frac{R + \theta\Pi - c\theta}{R - c\theta} \right)^{-\eta}} \quad (2.46)$$

On évalue le comportement de l'effort d'arrosage θ en fonction du couple de variables (p, π_{SI}) en supposant que l'agriculteur fait face d'abord à une sécheresse d'intensité maxi-

male. Le comportement de θ et du degré d'aversion au risque η de l'exploitant est apprécié en fonction des valeurs du couple de variables (p, π_{SI}) dépendant des conditions météorologiques pour des valeurs fixées du revenu initial R et du coût marginal c de l'effort investi.

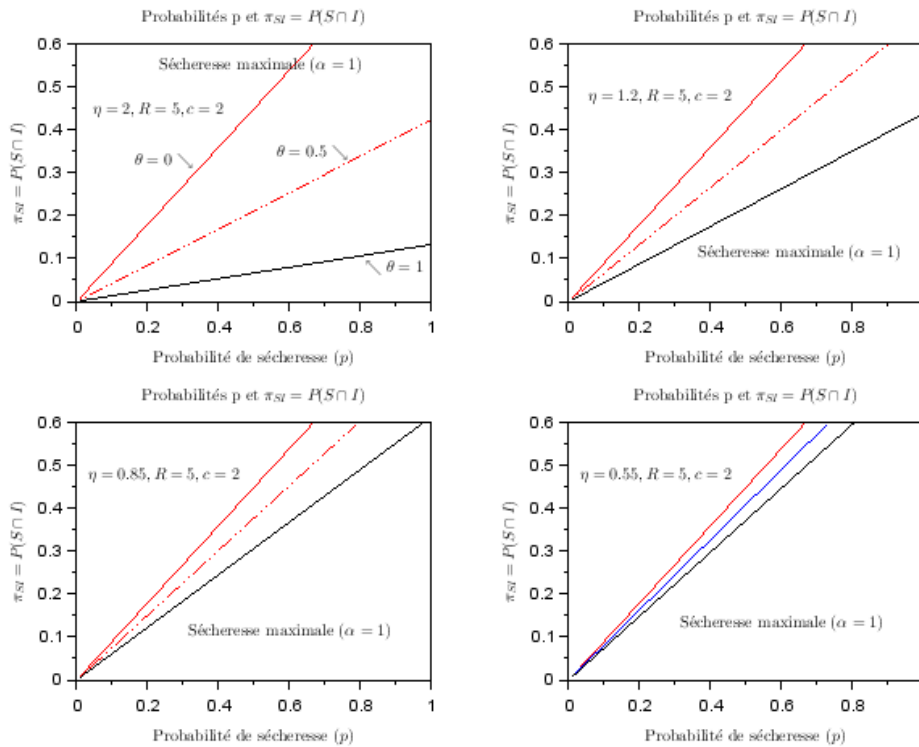
On représente graphiquement (figure 2.4.1) la fonction déterminée en (2.46) en fixant d'abord le niveau de revenu $R = 10$ et $c = 2$. Pour des valeurs typiques de $\eta \in \{0.55, 0.85, 1.2, 2\}$ (Holt et Laury, 2002) et $\theta \in \{0, 0.5, 1\}$, une augmentation de la probabilité qu'il se produise une sécheresse en faisant augmenter le niveau de risque sécheresse/inondation π_{SI} , induit un niveau d'effort d'arrosage moins important pour des niveaux de risques plus élevés. On peut remarquer qu'à probabilité de sécheresse égale, des niveaux de corrélation sécheresse/inondation plus élevés incitent à moins d'effort d'arrosage. Autrement dit, une augmentation de la probabilité de sécheresse induit une augmentation de la corrélation sécheresse/inondation ce qui a pour conséquence une réduction de l'effort d'arrosage. Cette réduction de l'effort d'arrosage est d'autant plus importante que l'agriculteur est averse au risque pour R et c fixés.

FIGURE 2.4.1 – Effort d'arrosage, aversion au risque et sécheresse maximale



Dans un deuxième temps, on divise le revenu initial R par deux en gardant les autres paramètres constants. Le comportement de θ et du degré d'aversion au risque η de l'exploitant en fonction du couple de variables (p, π_{SI}) se résume dans les graphiques suivants :

FIGURE 2.4.2 – Comportement de l'effort d'arrosage θ , du coefficient d'aversion au risque η suite à la variation de p et π_{SI}



Les analyses précédentes (figure 2.4.1) restent toujours valables. On peut faire remarquer que la réduction de 50% (figure 2.4.2) du revenu initial R , comparé à son niveau de départ, induit une diminution plus importante de l'effort d'arrosage. Cette réduction est d'autant plus importante que l'exploitant est averse au risque : d'autant plus faible est son revenu, d'autant plus importante est l'aversion de l'exploitant au risque (Mossin, 1968) et d'autant moins important est l'effort d'arrosage investi.

2.4.2 Revenu, revenu espéré, sécheresse et effort d'arrosage

D'après les résultats du modèle, l'activité d'arrosage est rentable et impacte positivement le revenu espéré dans la mesure où il existe une corrélation négative entre la survenue d'un épisode de sécheresse et celle d'une inondation succédant à la sécheresse.

Le revenu espéré défini à partir de l'équation (2.3) pour l'exploitant qui fait un effort maximal ($\theta = 1$) pour arroser sa parcelle est donné par :

$$E(R) = R + (1 - q)\Pi - pc \quad (2.47)$$

Dans le cas où *aucun effort* n'est fait ($\theta = 0$), le revenu espéré de l'exploitant est réduit à :

$$E(R) = R + [(1 - q) - \alpha(p - \pi_{SI})]\Pi. \quad (2.48)$$

Pour un niveau d'effort d'arrosage *intermédiaire* ($0 < \theta < 1$), le revenu espéré est donné par :

$$E(R) = R + (1 - q)\Pi - (p - \pi_{SI})(1 - \theta)\alpha\Pi - pc\theta. \quad (2.49)$$

A partir de cette dernière expression, l'impact de l'effort sur revenu espéré est tel que

$$\begin{aligned} \frac{\partial E(R)}{\partial \theta} &= (p - \pi_{SI})\alpha\Pi - pc \\ &= (\alpha\Pi - c)p - \alpha\Pi\pi_{SI}. \end{aligned}$$

Cet impact est positif que si

$$1 - \frac{c}{\alpha\Pi} > \frac{\pi_{SI}}{p}.$$

Ceci correspond exactement à la condition (2.23) garantissant un différentiel de revenu espéré positif quand il y a arrosage. D'un autre côté, l'impact de l'effort d'arrosage dépend du degré de vraisemblance de l'arrivée d'une sécheresse suivie par une inondation dévastatrice. On peut voir que

$$\frac{\partial}{\partial \pi_{SI}} \left(\frac{\partial E(R)}{\partial \theta} \right) = -\alpha\Pi < 0.$$

D'autant plus vraisemblable est l'arrivée d'une sécheresse suivie par une inondation dé-

vastatrice, d'autant moins important est l'effort investi : les perspectives de revenu espéré sont carrément mauvaises. Par contre, si on ne craint pas la survenue d'un épisode d'inondation, l'effort d'arrosage est bénéfique en situation de sécheresse car :

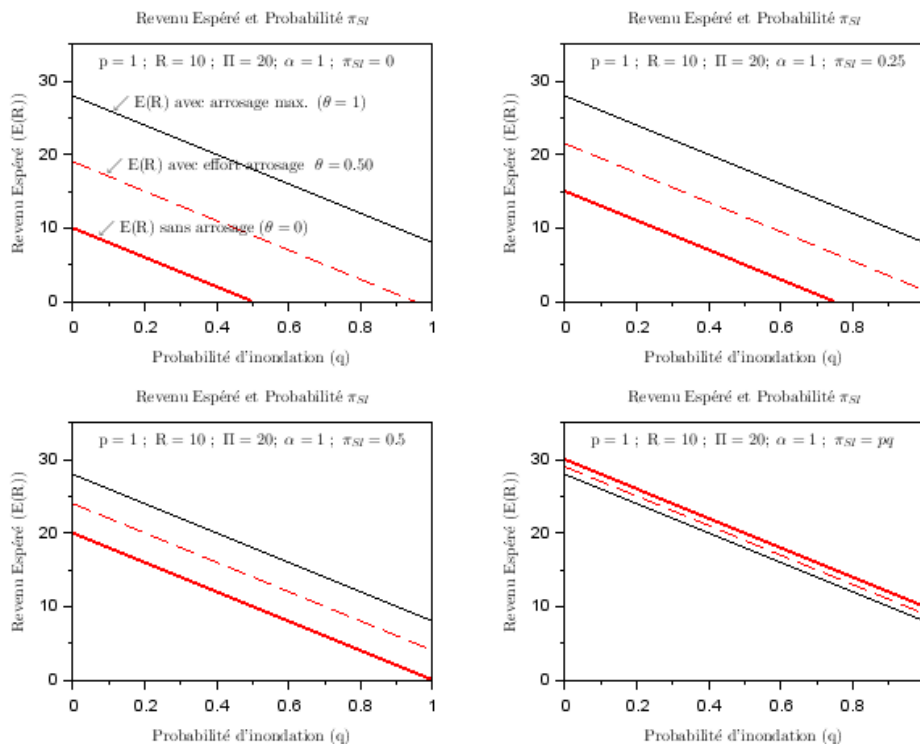
$$\frac{\partial}{\partial p} \left(\frac{\partial E(R)}{\partial \theta} \right) = \alpha \Pi - c$$

qui n'est autre que le profit net à condition que $\alpha \Pi > c$.

Par contre, on comprend que les perspectives de revenu espéré sont mauvaises d'autant plus que la survenue d'une inondation est vraisemblable. En effet,

$$\frac{\partial E(R)}{\partial q} = -\Pi < 0.$$

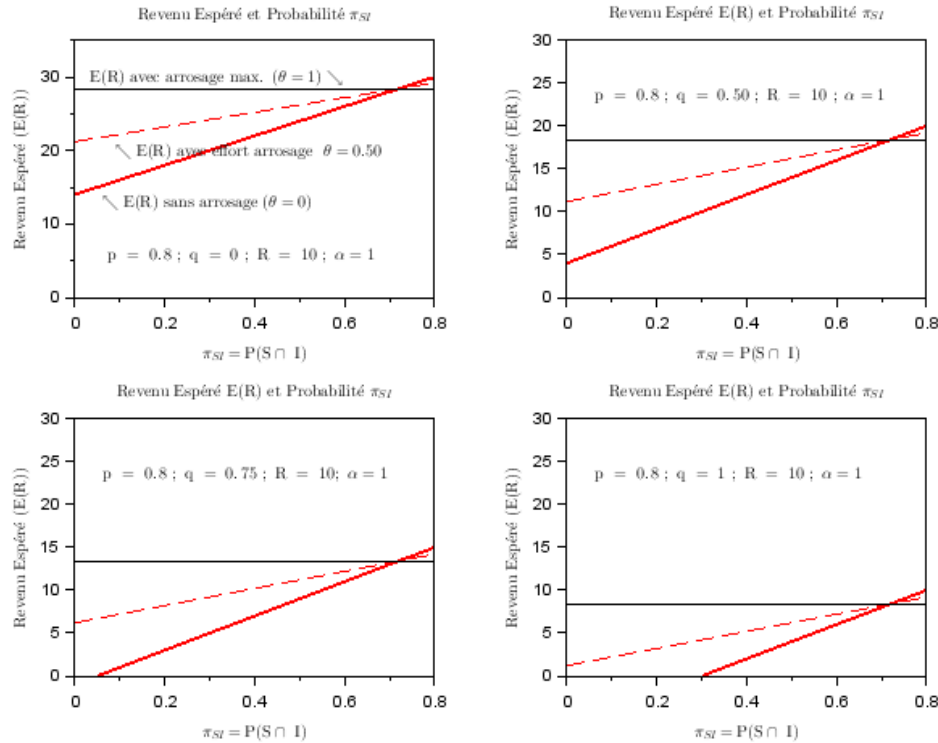
FIGURE 2.4.3 – Revenu espéré, Probabilité sécheresse et inondation et effort d'arrosage



Effectivement, le revenu espéré de l'exploitant s'associe négativement à la probabilité q qu'il survienne une inondation. En effet, quand $\pi_{SI} \rightarrow pq$ ($|\rho_{SI}| \rightarrow 0$), les perspectives de revenu espéré, diminuant avec l'augmentation de q , sont d'autant plus mauvaises en pré-

sence de zéro effort d'arrosage (courbe en rouge, figure 2.4.3) et d'autant moins mauvaises en présence d'un effort d'arrosage maximal (courbe en noir, figure 2.4.3).

FIGURE 2.4.4 – Revenu espéré, probabilité sécheresse et inondation et effort d'arrosage

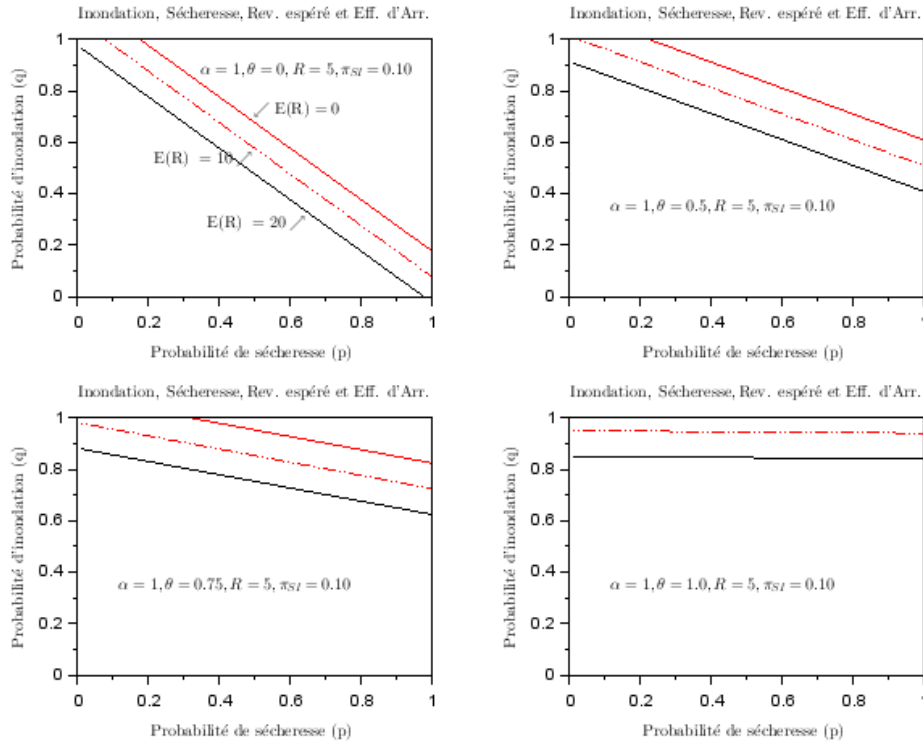


On peut comprendre qu'un risque élevé d'inondation, quand la sécheresse est sévère, impacte négativement l'effort d'arrosage et par voie de conséquence le niveau du revenu espéré. Cet impact est d'autant plus important que π_{SI} est proche de zéro et donc la corrélation (négative) entre risque de sécheresse et risque d'inondation est importante. D'où la nécessité que les petits exploitants qui consentent un effort d'arrosage soient couverts contre le risque d'inondation pour se prémunir contre des pertes potentielles de leur effort.

Les courbes de niveau évaluant comment se comporte le revenu espéré $E(R)$ en présence du couple (p, q) aident à mieux comprendre ces observations.

En gardant constants les autres paramètres et quand l'intensité de la sécheresse est maximale, si la survenue d'une inondation est de moins moins certaine, une augmentation de la probabilité de sécheresse p , en induisant une effort d'arrosage plus important (figure 2.4.5), fait augmenter les perspectives de revenu. Dans ce contexte, l'effort maximal

FIGURE 2.4.5 – Revenu espéré, probabilités sécheresse et inondation et effort d'arrosage



quand il est atteignable, est associé à des probabilités d'inondation moins élevées pour des probabilités de sécheresse plus élevées. D'autres graphiques non reportés ici indiquent que quand l'intensité de la sécheresse diminue, en maintenant à chaque fois constants les autres paramètres, un effort d'arrosage moins important est fait réduisant ainsi la zone d'effort d'arrosage. Par contre, quand la corrélation sécheresse/inondation est négative et élevée, la zone d'effort d'arrosage s'élargit.

2.4.3 Couverture d'assurance sur les inondations

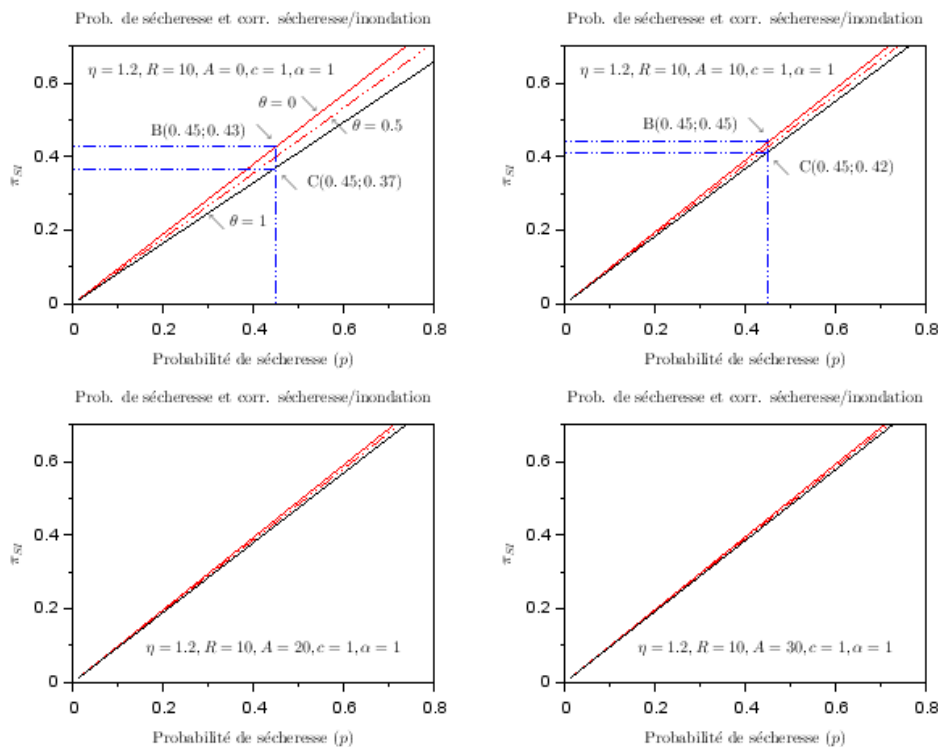
2.4.3.1 Couverture d'assurance exogène

Comme nous l'avons fait remarquer, l'agriculteur est incité à faire plus d'effort d'arrosage quand il a la garantie, en fonction de l'effort déployé, d'avoir une aide financière exogène A en cas d'inondation. En explicitant π_{SI} à partir de CPO (2.26), on obtient pour $\alpha = 1$ et $\theta \in [0, 1]$:

$$\pi_{SI} = \frac{p \left[\left(\frac{\Pi - c}{c} \right) \left(\frac{R + \theta \Pi - c\theta}{R - c\theta + A} \right)^{-\eta} \right]}{1 + \left(\frac{\alpha \Pi - c}{c} \right) \left(\frac{R + \theta \Pi - c\theta}{R - c\theta + A} \right)^{-\eta}}. \quad (2.50)$$

On peut voir qu'en analysant les graphiques de la figure 2.4.6, une sécheresse d'intensité maximale de plus en plus probable, en faisant augmenter π_{SI} , induit un effort d'arrosage de moins en moins important. Dans ce contexte, une augmentation de l'aide financière A incite l'exploitant à investir un certain effort d'arrosage. Le changement positif dans les coordonnées de points $B(\cdot)$ et $C(\cdot)$ quand l'aide financière A passe $A = 0$ à $A = 10$ témoigne de cet impact positif.

FIGURE 2.4.6 – Probabilité de sécheresse et probabilité conjointe sécheresse/inondation



L'impact des programmes d'assurance récolte aux Etats-Unis a été économétriquement testé par Skees et al. (1999). Il a pu établir qu'empiriquement, ces programmes d'assurance ont incité les agriculteurs à planter davantage, en particulier dans les régions les plus risquées.

2.4.3.2 Couverture d'assurance endogène

Couple optimal (A^*, θ^*) et courbes de niveaux (q, p) . Pour une prime ρA^* , un niveau d'assurance inondation rentable A^* optimal, corrélié à un niveau d'effort optimal θ^* , les courbes de niveaux (q, p) déterminées dans l'annexe (2.A.2.1) sont telles que

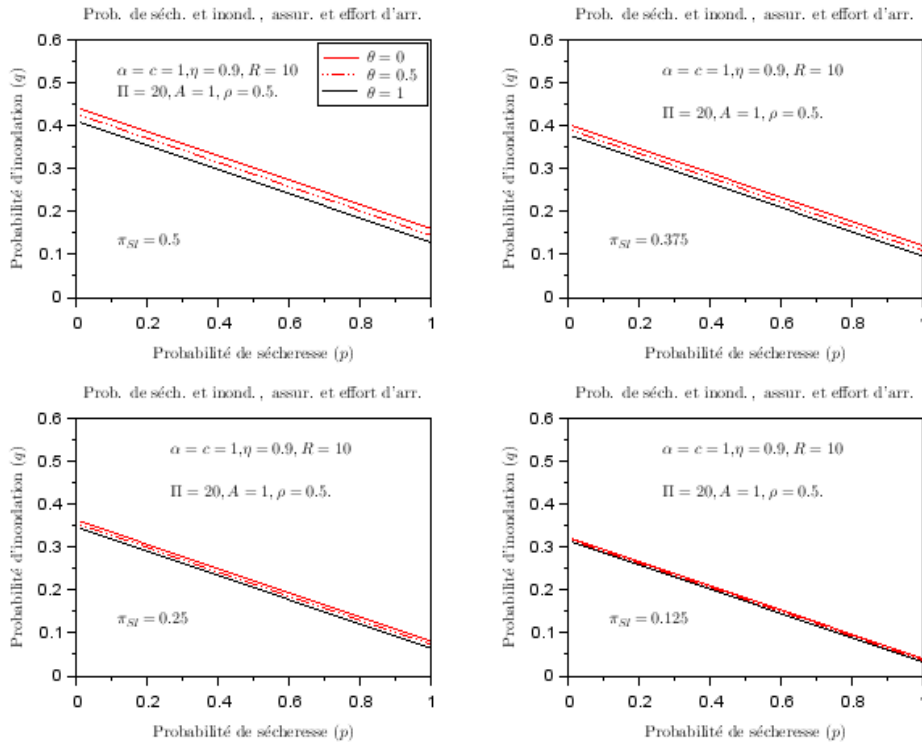
$$\begin{aligned}
q &= (1-p) \left(\frac{\rho}{D_q} \right) (R - \rho A + \Pi)^{-\eta} + [\rho (R - \rho A + \Pi)^{-\eta} \\
&\quad - \left(1 - \rho \left(\frac{\alpha \Pi}{\alpha \Pi - c} \right) \right) (R - \rho A - c\theta + A)^{-\eta} \\
&\quad + (1-\rho) (R - \rho A + A)^{-\eta}] \left(\frac{\pi_{SI}}{D_q} \right)
\end{aligned} \tag{2.51}$$

où

$$D_q = \rho (R - \rho A + \Pi)^{-\eta} + (1-\rho) (R - \rho A + A)^{-\eta}.$$

On trace et on analyse le comportement des courbes de niveaux (q, p) (fonction 2.51) pour $\rho \in \{1/4, 5/12, 7/12, 3/4\}$, $\pi_{SI} \in \{1/8, 1/4, 3/8, 1/2\}$ et des valeurs optimales $A^* = 1$ et $\theta^* \in \{0.0, 0.5, 1\}$ en fixant la valeur des autres paramètres

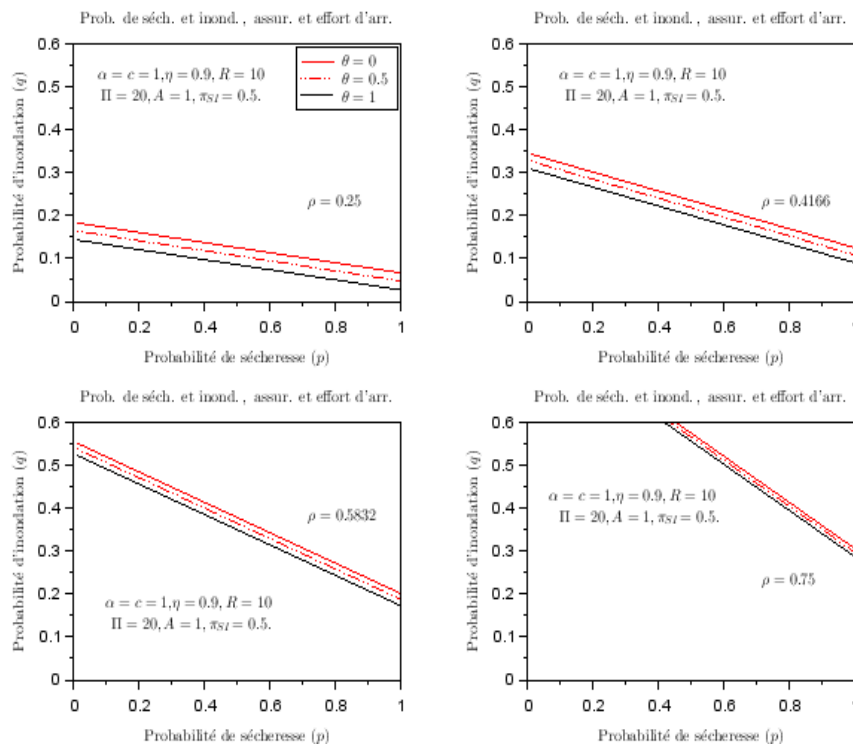
FIGURE 2.4.7 – Probabilités de sécheresse et d'inondation, assurance endogène et effort d'arrosage optimaux : $\pi_{SI} \in \{1/8, 1/4, 3/8, 1/2\}$.



Dans un premier temps, on fixe la valeur des autres paramètres et on considère $\pi_{SI} \in \{1/8, 1/4, 3/8, 1/2\}$. On peut faire les remarques suivantes (figure 2.4.7). Pour une probabilité d'un risque de sécheresse donnée, un niveau d'effort nul ($\theta = 0$), comparé

à un niveau d'effort maximal ($\theta = 1$), s'associe à une probabilité d'un risque d'inondation plus élevée. Celle-ci est d'autant plus élevée, comparée à une valeur plus faible de π_{SI} , que la corrélation positive sécheresse/inondation, induite par une augmentation de π_{SI} , est élevée. Car, que quelque soit le point $A_{is}(q_i, p_s)$ choisi sur les courbes de niveaux, $q_i p_s < \pi_{SI}$ et donc quand $\pi_{SI} \nearrow$, non seulement $\rho_{SI} > 0$ mais encore $\rho_{SI} \nearrow$.

FIGURE 2.4.8 – Probabilités de sécheresse et d'inondation, assurance endogène et effort d'arrosage optimaux : $\rho \in \{1/4, 5/12, 7/12, 3/4\}$



Dans un deuxième temps, on considère $\rho \in \{1/4, 5/12, 7/12, 3/4\}$ en fixant la valeur des autres paramètres (figure 2.4.8). Les remarques faites dans le cas de la figure 2.4.7 restent valables pour le taux de prime ρ à mesure que celui-ci augmente.

2.5 Les résultats du modèle eu égard à la littérature

Le développement et le potentiel de l'irrigation privée à petite échelle mise en place par les petits exploitants agricoles à faibles revenus ont attiré l'attention de beaucoup de chercheurs. De manière générale, les chercheurs ont pointé l'impact bénéfique de l'irrigation privée à petite échelle sur l'amélioration des revenus des exploitants en particulier,

le renforcement de la sécurité alimentaire et la réduction de la pauvreté en général. En Afrique sub-saharienne, le secteur des petites exploitations pratiquant l'irrigation privée à petite échelle offre des revenus importants à des millions de petits exploitants (Giordano et de Fraiture, 2014; Xie et al., 2014; De Fraiture et Giordano, 2014; Burney et Naylor, 2012). Tout ceci contribue à améliorer la sécurité alimentaire et réduire la pauvreté. Dillon (2011) a estimé un modèle d'appariement des coefficients de propension pour évaluer empiriquement l'effet de l'irrigation à petite et à grande échelles sur la production agricole et le revenu des agriculteurs. Il a pu établir que les secteurs pratiquant de l'irrigation à petite échelle bénéficient d'une amélioration plus substantielle de leur production et de leurs revenus, la consommation percapita étant plus importante dans les secteurs faisant usage de l'irrigation à grande échelle. D'un autre côté, Mango et al. (2018) a évalué économétriquement l'effet positif et significatif de l'adoption d'un système d'irrigation privée à petite échelle sur l'amélioration des revenus des exploitants du Triangle de Chinyanja en Afrique du Sud. Une analyse plus détaillée et une bibliographie plus exhaustive des expériences d'irrigation à petite échelle réalisées et leur impact peuvent être trouvées dans l'ouvrage publié par Merrey et Lefore (2018).

Les travaux précédemment cités sont des recherches descriptives et/ou empiriques. D'autres recherches surtout théoriques ont modélisé, dans un contexte de changement climatique, la problématique de l'assurabilité des risques climatiques. Nelson et Loehman (1987) ont développé un modèle dans lequel ils ont établi que théoriquement un optimum de Pareto de premier rang peut être atteint dans le cadre d'une couverture complète d'assurance agricole. Dans la pratique, des solutions de second rang telles par exemple des subsides de l'état peuvent être envisagées bien que non forcément nécessaires. Dans cet article, les différentes problématiques classiques liées à la théorie de l'assurance telles l'antiselection, l'hasard moral, l'asymétrie d'information et l'absence d'incentives adéquats ont été abordées. Les problèmes liés aux risques climatiques auxquels fait face le secteur agricole dans un contexte de changement climatique n'ont pas été systématiquement traités. Chambers (1989) a étendu le modèle de Nelson et Loehman (1987) en y incorporant un vecteur aléatoire prenant en compte les risques climatiques et d'autres sources de risque

impactant la production et le revenu des exploitants. Sous ces conditions, ces chercheurs ont analysé les différentes problématiques classiques liées à la théorie de l'assurance. Contrairement à Nelson et Loehman (1987), Chambers (1989) a prouvé qu'une neutralité au risque des agriculteurs ne garantit pas une couverture complète d'assurance pareto-optimale. Cependant, l'impact des variables climatiques n'a pas été explicitement analysé dans le cadre de ce modèle.

Le modèle que nous développons, bien que différent, s'inscrit en partie dans la démarche du modèle élaboré par Mahul (2001). Ce dernier a défini une fonction de production stochastique soumise à deux risques stochastiquement covariants : un risque climatique et un risque agrégé de production non assurable provenant de l'input utilisé. Dans son modèle, l'impact du risque climatique sur la production a deux origines : un impact lié à l'input utilisé et un impact résultant de la corrélation entre le risque climatique et le risque agrégé de production. Dans ce cas, le risque climatique est supposé assurable s'il n'est pas covariant. L'auteur a établi qu'un contrat d'assurance optimal existe pour un certain indice climatique conditionnellement à la dépendance entre le risque climatique assurable, le risque agrégé de production non assurable et l'attitude de l'exploitant à l'égard du risque. En fonction des hypothèses du modèle, il existe un indice climatique déclencheur garantissant une indemnité optimale. Si l'exploitant est averse au risque (dominance stochastique du premier ordre) et de plus prudent (dominance stochastique du second ordre), alors le risque agrégé de production devient plus important avec la diminution du risque climatique. Evidemment, les risques climatiques spécifiques de sécheresse et d'inondation n'ont pas été traités.

Le modèle que nous développons emprunte en partie la démarche suivie par Mahul (2001) avec un objectif plus spécifique : évaluer l'impact d'un effort d'arrosage et d'une assurance inondation en présence de risques corrélés de sécheresse et d'inondation.

2.6 Synthèse des résultats et conclusion

L'objectif essentiel de cet essai est d'évaluer théoriquement les potentialités et les limites d'un effort d'arrosage privé à petite échelle investi par un petit exploitant ⁶agricole dans un contexte climatique de risques de sécheresse et d'inondation. Le modèle élaboré à cet effet est monopériodique. Il décrit le comportement d'un petit agriculteur qui n'a d'autres choix pour s'en sortir que d'investir un certain effort d'arrosage privé. On a supposé qu'une intervention minimale de l'état en cas de sinistre est faisable surtout à posteriori. Les deux risques climatiques pris en compte - la sécheresse et l'inondation - ont permis de définir quatre états de la nature possibles : une situation de sécheresse immédiatement suivie par une inondation dévastatrice, un épisode de sécheresse non suivi par une inondation, l'arrivée tous azimuts d'une inondation sans précedence de sécheresse et une situation naturelle clémente - ni sécheresse ni inondation. De ce contexte, en adoptant une fonction d'utilité croissante, au moins deux fois dérivable et concave, l'utilité et le revenu espérés de l'exploitant ont été définis.

Les conditions de premier ordre de l'utilité espérée de l'exploitant par rapport à l'effort d'arrosage investi s'expriment, entre autres variables, en fonction de la distribution conjointe des risques sécheresse et inondation. Ceteris paribus, on a pu expliciter dans le cadre du modèle que si à chaque fois qu'il y a une sécheresse, il y a presque toujours aussi une inondation, dénotant ainsi une association positive entre ces deux états de la nature, l'exploitant averse aux risques positivement corrélés n'arrose pas. L'agriculteur n'est non plus incité à arroser quand les risques de sécheresse et d'inondation sont indépendants. L'agriculteur arrose sa parcelle seulement quand il est sûr de récupérer le fruit de ses efforts c'est-à-dire quand les années de sécheresse sont aussi celles où il n'y a pratiquement jamais d'inondation. Nous avons pu établir formellement que l'aversion aux risques de sécheresse et d'inondation corrélés positivement décourage l'exploitant à investir de moins en moins d'effort d'arrosage à mesure que la probabilité qu'un épisode de sécheresse suivi

6. On a voulu commencer par une modélisation des comportements individuels étant donné l'importance du nombre d'exploitants individuels en Haïti où près de 74% des exploitations ont une superficie de moins d'1 carreau (moins de 1ha29) et où un peu plus de 98% de ces exploitations sont gérées directement par les exploitants. On envisage par la suite de réévaluer le modèle en considérant des comportements de groupes plutôt qu'individuels.

de la survenue d'une inondation augmente.

En adoptant l'hypothèse que l'aversion au risque décroît avec le revenu, une amélioration des revenus de l'exploitant, averse au risque et aux risques de sécheresse et d'inondation positivement corrélés, constitue donc un élément qui l'incite à investir davantage d'effort d'arrosage dans une situation de sécheresse telle que les risques de sécheresse et d'inondation sont non positivement corrélés.

D'un autre côté, le risque élevé d'inondation ou une association positive entre les risques de sécheresse et d'inondation, en décourageant le moindre effort d'arrosage, impacte négativement les perspectives de revenu espéré. Quand l'intensité de la sécheresse est maximale, les illustrations effectuées indiquent que si la survenue d'une inondation est de moins en moins certaine, une augmentation de la probabilité de sécheresse p , en induisant un effort d'arrosage plus important, fait augmenter les perspectives de revenu de l'exploitant. Quand la corrélation sécheresse/inondation est négative et élevée, la zone d'effort d'arrosage s'élargit⁷. D'où la nécessité que les petits exploitants qui consentent un effort d'arrosage soient couverts contre le risque d'inondation pour se prémunir contre des pertes potentielles de leur effort.

L'impact négatif sur l'effort d'arrosage et le revenu espéré, d'un risque élevé d'inondation ou d'une association positive entre les risques de sécheresse et d'inondation, suggère la mise en place d'une protection des exploitants contre le risque d'inondation. Deux cas de figure sont envisagés. Une aide financière exogène pré ou post-inondation et la contractualisation d'une police d'assurance inondation. Une aide financière incite l'exploitant à investir davantage d'effort d'arrosage. Même en imposant une taxe A à des fins de compensation en absence d'inondation, l'exploitant est toujours incité à fournir, avec l'octroi d'une aide financière, un effort d'arrosage en cas de sécheresse toutes les fois que la corrélation sécheresse-inondation est négative. La contractualisation d'une police d'assurance endogène

7. Dans le cadre de cette modélisation, l'ensemble des états de la nature considérés est de type polytomique. Dans ce cas, ne peut être réalisée l'évaluation de situations intermédiaires telles par exemple l'impact de l'humidité des sols sur la réduction de l'intensité des inondations. L'humidité des sols constitue un facteur qui contribue à *l'interception du ruissellement intense* et du même coup réduit le "volume d'eau conduisant à l'inondation" (Breil, 2018). En effet, l'humidité des sols due à l'arrosage peut limiter les dommages pour certaines catégories d'inondation mineures et donc permettre de diminuer le coefficient de réduction de profit en cas d'inondation.

inondation impacte tout aussi positivement l'effort d'arrosage investi. Une couverture complète d'assurance inondation n'est pas possible du fait de la présence de risques résiduels ou inhérents liés aux coûts d'arrosage.

Dans ce cas précis, deux résultats essentiels sont établis. Dans un premier temps, il est démontré qu'en présence de risques corrélés de sécheresse et d'inondation, l'impossibilité d'avoir une assurance inondation rentable et où l'exploitant ne fait jamais l'effort d'arroser, sont des situations sans espoir. Ces situations correspondent naturellement à celles où il est de plus en plus probable qu'il survienne une inondation, quand il apparait de moins en moins probable qu'il s'installe une sécheresse, décourageant ainsi le moindre effort d'arrosage.. Dans un second temps, on établit qu'il n'est pas possible de trouver un niveau d'effort θ et une police d'assurance inondation A optimaux qui permettent d'éliminer tout risque bien qu'il existe un couple (θ^*, A^*) qui maximise l'utilité espérée de l'exploitant.

Bibliographie

- ANTON, J. et AL. (2013b). Agricultural risk management policies under climate uncertainty. *Global Environmental Change*, 23(6):1726–1736.
- BAR-SHIRA, Z. A. et AL. (1997). Estimation of farmers' risk attitude : an econometric approach. *Agricultural Economics*, 17:211–222.
- BARRETT, C. B. (2011). Covariate catastrophic risk management in the developing world : Discussion, forthcoming. *American Journal of Agricultural Economics*.
- BREIL, P. (2018). Intérêt des zones humides pour la prévention des inondations par ruissellement. *Sciences Eaux Territoires*, pages 62–65.
- BRIYS, E. et SCHLESINGER, H. (1990). Risk aversion and the propensities for self insurance and self protection. *Southern economic Journal*, 57(2):458–467.
- BURNEY, J. A. et NAYLOR, R. L. (2012). Smallholder irrigation as a poverty alleviation tool in sub saharan africa. *World Development*, 40(1):110–123.
- CARTER, R. C. (1989). The development of small scale irrigation in sub Saharan Africa. *Public Administration and Development*, 9:543–555.
- CHAMBERS, R. G. (1989). Insurability and moral hazard in agricultural insurance markets. *American Journal of Agricultural Economics*, 71(3):604–616.
- CHARPENTIER, A. (2008). Insurability of climate risks. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 33(1):91–109.
- CII (2009). Market failure and climate change (chapter 5) in "coping with climate change risks and opportunities for insurers". Rapport technique, The Chartered Insurance Institute (CII).
- DAI, B. ; Ding, S. et WAHBA, G. (2013). Multivariate bernoulli distribution. *Bernoulli*, 19(4):1465–1483.

- DE FRAITURE, C. et GIORDANO, M. (2014). Small private irrigation : A thriving but overlooked sector. *Agricultural Water Management*, 131:167–174.
- DENUIT, M. M.; Eeckhoudt, L. et REY, B. (2010). Some consequences of correlation aversion in decision science. *Annals of Operations Research*, 176(1):259–269.
- DILLON, A. (2011). Do differences in the scale of irrigation projects generate different impacts on poverty and production ? *Journal of Agricultural Economics*, 62(2):474–492.
- DIONNE, G. et ECKHOUDT, L. (1985). Self insurance, self protection and increased risk aversion. *Economics Letters*, 17(1-2):39–42.
- DOHERTY, N. (1997). Insurance markets and climate change. *Geneva Papers on Risk and Insurance. Issues and Practice*, pages 223–237.
- EPSTEIN, L. G. et TANNY, S. M. (1980). Increasing generalized correlation : a definition and some economic consequences. *Canadian Journal of Economics*, pages 16–34.
- FANG, W. et AL. (2019). Copulas based risk analysis for inter seasonal combinations of wet and dry conditions under a changing climate. *International Journal of Climatology*, 39(4):2005–2021.
- FEDER, G. ; Just, R. E. et ZILBERMAN, D. (1985). Adoption of agricultural innovations in developing countries : A survey. *Economic development and cultural change*, 33(2): 255–298.
- FINKELSHTAIN, I. ; Kella, O. et SCARSINI, M. (1999). On risk aversion with two risks. *Journal of Mathematical Economics*, 31(2):239–250.
- GIEC et PACHAURI, A. R. (2014). Changements climatiques 2014 : Rapport de synthèse. contribution des groupes de travail i, ii et iii . cinquième rapport d'évaluation. Rapport technique, Groupe D'experts Intergouvernemental sur L'évolution du Climat (GIEC), 60-79.
- GIORDANO, M. et de FRAITURE, C. (2014). Small private irrigation : Enhancing benefits and managing trade offs. *Agricultural Water Management*, 131:175–182.

- GLAUBER, J. W.; Keith, J. C. et PETER, J. B. (2002). Crop insurance, disaster assistance and the role of the federal government in providing catastrophic risk protection. *Agricultural Finance Review*.
- GOLLIER, C. (1996). Optimum insurance of approximate losses. *The Journal of Risk and Insurance*, 63(3):369.
- HAZELL, P. B. R. (1992). The appropriate role of agricultural insurance in developing countries. *Journal of International Development*, 4(6):567–581.
- HOLT, C. A. et LAURY, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 57(2):458–467.
- INADA, K. I. (1963). On a two sector model of economic growth : Comments and a generalization. *The Review of Economic Studies*, 30(2):119–127.
- JAFFEE, D. M. et RUSSELL, T. (1997). Catastrophe insurance, capital markets and uninsurable risks. *Journal of Risk and Insurance*, pages 205–230.
- JEHLE, G. A. et RENY, P. J. (2011). *Advanced Microeconomic Theory. Third edition*. Prentice Hall (Financial Times), Pearson.
- KILROY, G. (2015). A review of the biological impacts of climate change in three hotspot regions in africa and asia. *Regional Environmental Change*, 15(5):771–782.
- LUDWIG, F. et AL. (2007). Climate change impacts on developing countries. Rapport technique, EU accountability and European Parliament.
- MAHUL, O. (2001). Optimal insurance against climatic experience. *American Journal of Agricultural Economics*, 83(3):593–604.
- MAHUL, O. et STUTLEY, C. J. (2010). Government support to agricultural insurance : challenges and options for developing countries. Rapport technique, World Bank.
- MANGO, N. et AL. (2018). Adoption of small scale irrigation farming as a climate smart agriculture practice and its influence on household income in the chinyanja triangle, southern africa. *Land*, 7(2):49.

- MENDELSON, R. O. et DINAR, A. (2009). *Climate change and agriculture : an economic analysis of global impacts, adaptation and distributional effects*. Edward Elgar Publishing.
- MERREY, D. J. et LEFORE, N. (2018). Improving the availability and effectiveness of rural and micro finance for small scale irrigation in sub saharan africa : a review of lessons learned. *International Water Management Institute*, 185.
- MIRANDA, M. J. et GLAUBER, J. W. (1997). Systemic risk, reinsurance and the failure of crop insurance markets. *American journal of agricultural economics*, 79(1):206–215.
- MOSSIN, J. (1968). Aspects of rational insurance purchasing. *Journal of political economy*, 76(4):553–568.
- NATH, P. K. et BEHERA, B. (2011). A critical review of impact of and adaptation to climate change in developed and developing economies. *Environment, development and sustainability*, 13(1):141–162.
- NELSON, C. H. et LOEHMAN, E. T. (1987). Further toward a theory of agricultural insurance. *American Journal of Agricultural Economics*, 69(3):523–531.
- PALACIOS-HUERTA, I. et SERRANO, R. (2006). Rejecting small gambles under expected utility. *Economics Letters*, 91(2):250–259.
- PRATT, J. W. (1964). Risk aversion in the large and in the small. *Econometrica*, 32(1-2):122–136.
- RAI, P. K. et RAI, P. K. (2013). Environmental and socio economic impacts of global climate change : An overview on mitigation approaches. *Environmental Skeptics and Critics*, 2(4):126–148.
- RILEY JR, W. B. et CHOW, K. V. (1992). Asset allocation and individual risk aversion. *Financial Analysts Journal*, 48(6):32–37.
- SKEES, J. R. et REED, M. R. (1986). Rate making for farm level crop insurance : Implications for adverse selection. *American Journal of Agricultural Economics*, 68(3):653–659.

- SKEES, K. K. et AL. (1999). The potential influence of risk management programs on cropping decisions at the extensive margin. Mémoire de D.E.A., University of Kentucky.
- SURMINSKI, S. (2013). The role of insurance in reducing direct risk : the case of flood insurance. *International Review of Environmental and Resource Economics*, 7(3-4):241–278.
- SZPIRO, G. G. (1983). The hypotheses of absolute and relative risk aversion : An empirical study using cross section data. *Geneva Papers on Risk and Insurance*, pages 336–349.
- TILLIE, P. et AL. (2019). Impacts ex-ante de la Petite Irrigation au Niger. Analyse des effets micro-économiques à laide dun modèle de ménage agricole. *Joint Research Centre, Seville site*.
- WAKKER, P. P. (2008). Explaining the characteristics of the power (crra) utility family. *Health economics*, 17(12):1329–1344.
- WU, S. ; Zhang, X. et SHE, D. (2017). Joint pattern of seasonal hydrological droughts and floods alternation in china's huai river basin using the multivariate l-moments. *Journal of Earth System Science*, 126(4):1–17.
- XIE, H. et AL. (2014). Estimating the potential for expanding smallholder irrigation in sub saharan africa. *Agricultural Water Management*, 131:183–193.

Appendix

2.A Annexes

2.A.1 Résultats théoriques

2.A.1.1 Impact formel de π_{SI} sur l'effort d'arrosage optimal

En supposant l'existence d'une solution intérieure θ^* pour la *CPO* obtenue à l'équation (), on prend la dérivée par rapport à π_{SI} :

$$\begin{aligned} & \frac{d}{d\pi_{SI}} \{-\pi_{SI} c v' (R - c\theta^*) \\ & + (p - \pi_{SI}) [\alpha \Pi - c] v' (R + [\alpha\theta^* + (1 - \alpha)] \Pi - c\theta^*)\} = 0 \end{aligned}$$

Par suite

$$\begin{aligned} & -c v' (R - c\theta^*) + c^2 \pi_{SI} v'' (R - c\theta^*) \frac{d\theta^*}{d\pi_{SI}} \\ & + \{(\alpha \Pi - c) [-v' (R + (\alpha\theta^* + (1 - \alpha)) \Pi - c\theta^*) \\ & + (p - \pi_{SI}) v'' (R + [\alpha\theta^* + (1 - \alpha)] \Pi - c\theta^*) \left(\alpha \Pi \frac{d\theta^*}{d\pi_{SI}} - c \frac{d\theta^*}{d\pi_{SI}} \right) \right] \} = 0. \end{aligned}$$

En réarrangeant et en factorisant, on obtient :

$$\begin{aligned} & \frac{d\theta^*}{d\pi_{SI}} \{(\alpha \Pi - c)^2 (p - \pi_{SI}) v'' (R + [\alpha\theta^* + (1 - \alpha)] \Pi - c\theta^*) + c^2 \pi_{SI} v'' (R - c\theta^*)\} = \\ & [(\alpha \Pi - c) v' (R + (\alpha\theta^* + (1 - \alpha)) \Pi - c\theta^*) + c v' (R - c\theta^*)] \end{aligned}$$

de sorte que

$$\frac{d\theta^*}{d\pi_{SI}} = \frac{(\alpha\Pi - c)v'(R + (\alpha\theta^* + (1 - \alpha))\Pi - c\theta^*) + cv'(R - c\theta^*)}{D_1 + D_2} < 0 \quad (2.52)$$

où

$$D_1 = (\alpha\Pi - c)^2 (p - \pi_{SI}) v''(R + [\alpha\theta^* + (1 - \alpha)]\Pi - c\theta^*)$$

$$D_2 = c^2 \pi_{SI} v''(R - c\theta^*)$$

2.A.1.2 Impact du revenu sur l'effort d'arrosage

En dérivant la *CPO* (2.5) par rapport à R , on obtient ($0 < \theta < 1$)

$$\begin{aligned} 0 &= -\pi_{SI}c \left(1 - c \frac{d\theta}{dR}\right) v''(R - c\theta) \\ &+ (p - \pi_{SI}) [\alpha\Pi - c] \left(1 + (\alpha\Pi - c) \frac{d\theta}{dR}\right) v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta). \end{aligned}$$

En réarrangeant, en factorisant et en supposant

$$\pi_{SI}c^2 v''(R - c\theta) \neq (p - \pi_{SI}) (\alpha\Pi - c)^2 v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta),$$

on obtient :

$$\begin{aligned} \frac{d\theta}{dR} &= \frac{\pi_{SI}c v''(R - c\theta) - (p - \pi_{SI}) [\alpha\Pi - c] v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{[\pi_{SI}c^2 v''(R - c\theta) + (p - \pi_{SI}) (\alpha\Pi - c)^2 v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)]} \\ &= \frac{1}{c} \left\{ \frac{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) - \left(\frac{\alpha\Pi - c}{c}\right) [v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) / v''(R - c\theta)]}{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) + \left(\frac{\alpha\Pi - c}{c}\right)^2 [v''(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) / v''(R - c\theta)]} \right\} \\ &= \frac{1}{c} \left\{ \frac{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) - \left(\frac{\alpha\Pi - c}{c}\right) \left[\frac{v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{v'(R - c\theta)}\right] \left[\frac{a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta)}\right]}{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) + \left(\frac{\alpha\Pi - c}{c}\right)^2 \left[\frac{v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{v'(R - c\theta)}\right] \left[\frac{a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta)}\right]} \right\} \\ &= \frac{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) - \left[\frac{\pi_{SI}}{p - \pi_{SI}}\right] \left[\frac{a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta)}\right]}{\left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) + \left(\frac{\pi_{SI}}{p - \pi_{SI}}\right) (\alpha\Pi - c) \left[\frac{a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta)}\right]} \end{aligned} \quad (2.53)$$

$$= \frac{a_{RA}(R - c\theta) - a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}{a_{RA}(R - c\theta) + (\alpha\Pi - c) a_{RA}(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)}, \quad (2.54)$$

où le passage à l'équation (2.53) se fait en utilisant la condition (2.9) en se rappelant que $a_{RA}(R) = -u''(R)/u'(R)$.

De manière similiaire,

$$\frac{\pi_{SI}}{p - \pi_{SI}} = \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \left(\frac{\alpha\Pi - c}{c} \right)$$

et donc

$$\begin{aligned} \frac{d}{dR} \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) &= \left(\frac{\alpha\Pi - c}{c} \right) \frac{d}{dR} \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \\ &= \left(\frac{v''(R + (1 - \alpha)\Pi)v'(R) - v''(R)v'(R + (1 - \alpha)\Pi)}{[v'(R)]^2} \right) \\ &\quad \times \left(\frac{\alpha\Pi - c}{c} \right) \\ &= [a_{RA}(R) - a_{RA}(R + (1 - \alpha)\Pi)] \left(\frac{\alpha\Pi - c}{c} \right) \\ &\quad \times \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \\ &= [a_{RA}(R) - a_{RA}(R + (1 - \alpha)\Pi)] \left(\frac{\pi_{SI}}{p - \pi_{SI}} \right) \end{aligned} \quad (2.55)$$

2.A.1.3 Impact d'une assurance exogène sur les inondations

Quand l'exploitant es susceptible de recevoir une aide financière A en cas d'inondation, son utilité espérée s'écrit :

$$\begin{aligned} E(V) &= \pi_{SI}v(R - c\theta + A) + (q - \pi_{SI})v(R + A) \\ &\quad + (p - \pi_{SI})v(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &\quad + (1 - p - q + \pi_{SI})v(R + \Pi). \end{aligned} \quad (2.56)$$

L'effort optimal qui maximise l'utilité de l'exploitant dans le cas où il reçoit une financière A en cas d'inondation est donnée par la condition du premier ordre en θ :

$$\begin{aligned} 0 &= -\pi_{SI}cv'(R - c\theta + A) \\ &\quad + (p - \pi_{SI})[\alpha\Pi - c]v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta). \end{aligned} \quad (2.57)$$

Si on dérive cette dernière équation par rapport à A , on obtient :

$$0 = \left(c \frac{d\theta}{dA} - 1 \right) \pi_{SI} c v'' (R - c\theta + A) \\ + (\alpha\Pi - c)^2 \frac{d\theta}{dA} (p - \pi_{SI}) v'' (R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta).$$

soit

$$\pi_{SI} c v'' (R - c\theta + A) = \frac{d\theta}{dA} \left\{ \pi_{SI} c^2 v'' (R - c\theta + A) \right. \\ \left. + (p - \pi_{SI}) (\alpha\Pi - c)^2 v'' (R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \right\}.$$

Ainsi, en notant

$$v''_A(\cdot) = v'' (R - c\theta + A) \\ v''_{\bar{A}}(\cdot) = v'' (R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta)$$

obtient finalement

$$\frac{d\theta}{dA} = \frac{\pi_{SI} c v''_A(\cdot)}{\pi_{SI} c^2 v''_A(\cdot) + (p - \pi_{SI}) (\alpha\Pi - c)^2 v''_{\bar{A}}(\cdot)}. \quad (2.58)$$

Comme $v''(\cdot) < 0$, donc $(d\theta/dA) > 0$.

Impact d'une taxe sur l'effort d'arrosage. La *CPO* (2.30) par rapport à θ est telle que

$$\frac{dE(V)}{d\theta} = -\pi_{SI} c v' (R - c\theta + A) \\ + (p - \pi_{SI}) (\alpha\Pi - c) v' \{R + [\alpha\theta + (1 - \alpha) - \tau_A] \Pi - c\theta\}.$$

En supposant que $\theta = \theta(\tau_A)$, l'impact de cette taxe à des fins de compensation sur l'effort optimal d'arrosage est telle que

$$\begin{aligned} 0 &= c^2 \pi_{SI} c v'' (R - c\theta + A) \frac{d\theta}{d\tau_A} \\ &+ \{ (p - \pi_{SI}) (\alpha \Pi - c) v'' [R + (\alpha\theta + (1 - \alpha) - \tau_A) \Pi - c\theta] \\ &\times \left[\left(\alpha \frac{d\theta}{d\tau_A} - 1 \right) \Pi - c \frac{d\theta}{d\tau_A} \right] \} \end{aligned}$$

Soit

$$\begin{aligned} 0 &= c^2 \pi_{SI} c v'' (R - c\theta + A) \frac{d\theta}{d\tau_A} \\ &+ \{ (p - \pi_{SI}) (\alpha \Pi - c) v'' [R + (\alpha\theta + (1 - \alpha) - \tau_A) \Pi - c\theta] \\ &\times \left[\frac{d\theta}{d\tau_A} (\alpha \Pi - c) - \Pi \right] \}. \end{aligned}$$

Par suite, en notant

$$\begin{aligned} v''_{\tau}(\cdot) &= v'' (R - c\theta + A), \\ v''_{\tau}(\cdot) &= v'' [R + (\alpha\theta + (1 - \alpha) - \tau_A) \Pi - c\theta], \end{aligned}$$

on obtient :

$$\left[c^2 \pi_{SI} c v''_{\tau}(\cdot) + (p - \pi_{SI}) (\alpha \Pi - c)^2 v''_{\tau}(\cdot) \right] \frac{d\theta}{d\tau_A} = (p - \pi_{SI}) (\alpha \Pi - c) \Pi v''_{\tau}(\cdot)$$

ce qui donne

$$\frac{d\theta}{d\tau_A} = \frac{(p - \pi_{SI}) (\alpha \Pi - c) \Pi v''_{\tau}(\cdot)}{c^2 \pi_{SI} c v''_{\tau}(\cdot) + (p - \pi_{SI}) (\alpha \Pi - c)^2 v''_{\tau}(\cdot)} > 0. \quad (2.59)$$

2.A.1.4 Impact de l'effort d'arrosage sur la couverture d'assurance endogène

En dérivant la *CPO* donnée à l'équation (2.34) par rapport à θ , on obtient :

$$\begin{aligned}
0 = & \pi_{SI} \left((1 - \rho) \frac{dA}{d\theta} - c \right) (1 - \rho) v'' (R - \rho A - c\theta + A) \\
& + (q - \pi_{SI}) \left((1 - \rho) \frac{dA}{d\theta} \right) (1 - \rho) v'' (R - \rho A + A) \\
& - (p - \pi_{SI}) \left(-\rho \frac{dA}{d\theta} + \alpha \Pi - c \right) \rho v'' (R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\
& + (1 - p - q + \pi_{SI}) \rho^2 \frac{dA}{d\theta} v'' (R - \rho A + \Pi), \tag{2.60}
\end{aligned}$$

la couverture d'assurance A à contracter dépendant tout aussi de θ . Par suite, en notant

$$\begin{aligned}
v''_{SI}(\cdot) &= v'' (R - \rho A + A) \\
v''_{0I}(\cdot) &= v'' (R - \rho A - c\theta + A) \\
v''_{S0}(\cdot) &= v'' (R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\
v''_{00}(\cdot) &= v'' (R - \rho A + \Pi),
\end{aligned}$$

on obtient :

$$\begin{aligned}
& \pi_{SI} c (1 - \rho) v''_{0I}(\cdot) + (p - \pi_{SI}) (\alpha \Pi - c) \rho v''_{S0}(\cdot) = \\
& \pi_{SI} \left((1 - \rho) \frac{dA}{d\theta} \right) (1 - \rho) v''_{0I}(\cdot) + (q - \pi_{SI}) \left((1 - \rho) \frac{dA}{d\theta} \right) (1 - \rho) v''_{SI}(\cdot) \\
& + (p - \pi_{SI}) \left(\rho^2 \frac{dA}{d\theta} \right) v''_{S0}(\cdot) + (1 - p - q + \pi_{SI}) \rho^2 v''_{00}(\cdot)
\end{aligned}$$

ou encore

$$\begin{aligned}
& \pi_{SI} c (1 - \rho) v''_{0I}(\cdot) + (p - \pi_{SI}) (\alpha \Pi - c) \rho v''_{S0}(\cdot) = \\
& \frac{dA}{d\theta} \{ (1 - \rho)^2 [\pi_{SI} v''_{0I}(\cdot) + (q - \pi_{SI}) v''_{SI}(\cdot)] \\
& + \rho^2 [(p - \pi_{SI}) v''_{S0}(\cdot) + (1 - p - q + \pi_{SI}) v''_{00}(\cdot)] \}.
\end{aligned}$$

En posant

$$\begin{aligned}\Delta_{\theta A} &= (1 - \rho)^2 [\pi_{SI} v''_{0I}(\cdot) + (q - \pi_{SI}) v''_{SI}(\cdot)] \\ &\quad + \rho^2 [(p - \pi_{SI}) v''_{S0}(\cdot) + (1 - p - q + \pi_{SI}) v''_{00}(\cdot)]\end{aligned}$$

on obtient finalement :

$$\frac{dA}{d\theta} = \frac{\pi_{SI}}{\Delta_{\theta A}} c (1 - \rho) v''_{0I}(\cdot) + \frac{p - \pi_{SI}}{\Delta_{\theta A}} (\alpha \Pi - c) \rho v''_{S0}(\cdot) \quad (2.61)$$

Etant donné que $v''(\cdot) < 0$, on n'en déduit que $\Delta_{\theta A} < 0$ et donc $(dA/d\theta) > 0$.

Aucune couverture d'assurance. En considérant la *CPO* (2.34) par rapport à A , la condition $(dE(V)/dA) < 0$ se traduit par :

$$\begin{aligned}\pi_{SI} (1 - \rho) v'(R - \rho A - c\theta + A) \\ + (q - \pi_{SI}) (1 - \rho) v'(R - \rho A + A) < \\ (p - \pi_{SI}) \rho v'(R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\ + (1 - p - q + \pi_{SI}) \rho v'(R - \rho A + \Pi).\end{aligned}$$

ou de manière équivalente

$$\begin{aligned}\pi_{SI} v'(R - \rho A - c\theta + A) \\ + (q - \pi_{SI}) v'(R - \rho A + A) < \\ \rho [(p - \pi_{SI}) v'(R - \rho A + (\alpha\theta + (1 - \alpha)) \Pi - c\theta) \\ + (1 - p - q + \pi_{SI}) v'(R - \rho A + \Pi) \\ + \pi_{SI} v'(R - \rho A - c\theta + A) \\ + (q - \pi_{SI}) v'(R - \rho A + A)]\end{aligned}$$

et donc

$$\begin{aligned}\rho &> \frac{\pi_{SI}v'(R - \rho A - c\theta + A) + (q - \pi_{SI})v'(R - \rho A + A)}{\Delta_\rho} \\ &= \frac{\pi_{SI}}{\Delta_\rho}v'(R - \rho A - c\theta + A) + \frac{q - \pi_{SI}}{\Delta_\rho}v'(R - \rho A + A)\end{aligned}$$

avec

$$\begin{aligned}\Delta_\rho &= (p - \pi_{SI})v'(R - \rho A + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &\quad + (1 - p - q + \pi_{SI})v'(R - \rho A + \Pi) \\ &\quad + \pi_{SI}v'(R - \rho A - c\theta + A) + (q - \pi_{SI})v'(R - \rho A + A).\end{aligned}$$

Ainsi, quand $(dE(V)/dA) < 0$ alors $A = 0$ et donc

$$\rho_{A=0} > \frac{\pi_{SI}}{\Delta'_\rho}v'(R - c\theta) + \frac{(q - \pi_{SI})}{\Delta'_\rho}v'(R) \quad (2.62)$$

où

$$\begin{aligned}\Delta'_\rho &= (p - \pi_{SI})v'(R + [\alpha\theta + (1 - \alpha)]\Pi - c\theta) \\ &\quad + (1 - p - q + \pi_{SI})v'(R + \Pi) \\ &\quad + \pi_{SI}v'(R - c\theta) + (q - \pi_{SI})v'(R).\end{aligned}$$

Assurance endogène, effort d'arrosage maximal et prime actuarielle. En supposant que $A = \Pi + c$, la *CPO* (2.38) devient :

$$\begin{aligned}
\frac{dE(V)}{dA} &= (1-q)\pi_{SI}v'(R-q(\Pi+c)+\Pi+c-c) \\
&\quad -q(p-\pi_{SI})v'(R-q(\Pi+c)+\Pi-c) \\
&\quad -q(1-p-q+\pi_{SI})v'(R-q(\Pi+c)+\Pi) \\
&\quad + (1-q)(q-\pi_{SI})v'(R-q(\Pi+c)+\Pi+c) \\
&= (1-q)\pi_{SI}v'(R+(1-q)\Pi-qc) \\
&\quad -q(p-\pi_{SI})v'(R+(1-q)\Pi-(1+q)c) \\
&\quad -q(1-p-q+\pi_{SI})v'(R+(1-q)\Pi-qc) \\
&\quad + (1-q)(q-\pi_{SI})v'(R+(1-q)(\Pi+c)).
\end{aligned}$$

Soit, en réarrangeant,

$$\begin{aligned}
\frac{dE(V)}{dA} &= q(p-\pi_{SI})[v'(R+(1-q)\Pi-qc) \\
&\quad -v'(R+(1-q)\Pi-(1+q)c)] \\
&\quad + (1-q)(q-\pi_{SI})[v'(R+(1-q)(\Pi+c)) \\
&\quad -v'(R+(1-q)\Pi-qc)] \\
&> 0.
\end{aligned} \tag{2.63}$$

car $v'(\cdot) > 0$ et donc

$$\begin{aligned}
v'[R+(1-q)\Pi+(1-q)c] &> v'[R+(1-q)\Pi-qc] \\
&> v'[R+(1-q)\Pi-(1+q)c]
\end{aligned}$$

Pas d'arrosage et pas d'assurance. La *CPO* (2.33) par rapport à θ est telle que

$$\begin{aligned} \frac{\partial E(V)}{\partial \theta} &= -\pi_{SI} c v'(R - \rho A - c\theta + A) \\ &\quad + (p - \pi_{SI}) [\alpha \Pi - c] v'(R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta). \end{aligned}$$

Quand il n'y a pas d'effort d'arrosage ($\theta = 0$) et d'assurance ($A = 0$) en situation d'une sécheresse d'intensité quelconque α ($0 < \alpha < 1$), alors l'impact de θ sur l'utilité de l'exploitant est tel que $\partial E(V) / \partial \theta < 0$. Par suite, de la *CPO* (2.33) par rapport à θ :

$$-\pi_{SI} c v'(R) + (p - \pi_{SI}) (\alpha \Pi - c) v'(R + (1 - \alpha) \Pi) < 0$$

soit

$$\left(\frac{c}{\alpha \Pi - c} \right) \pi_{SI} v'(R) > (p - \pi_{SI}) v'(R + (1 - \alpha) \Pi) \quad (2.64)$$

ce qui donne enfin

$$\frac{\pi_{SI}}{p - \pi_{SI}} > \left(\frac{v'(R + (1 - \alpha) \Pi)}{v'(R)} \right) \frac{\alpha \Pi - c}{c}. \quad (2.65)$$

Etant donné que la *CPO* 2.34 en A est telle que :

$$\begin{aligned} \frac{\partial E(V)}{\partial A} &= \pi_{SI} (1 - \rho) v'(R - \rho A - c\theta + A) \\ &\quad + (q - \pi_{SI}) (1 - \rho) v'(R - \rho A + A) \\ &\quad - (p - \pi_{SI}) \rho v'(R - \rho A + [\alpha\theta + (1 - \alpha)] \Pi - c\theta) \\ &\quad - (1 - p - q + \pi_{SI}) \rho v'(R - \rho A + \Pi). \end{aligned}$$

Quand $\theta = 0$ et $A = 0$, l'utilité de l'agriculteur est telle que $\partial E(V) / \partial A < 0$. Ce qui se traduit par (équation 2.34) :

$$\begin{aligned} &\pi_{SI} v'(R) - \pi_{SI} \rho v'(R) + (q - \pi_{SI}) v'(R) - \rho (q - \pi_{SI}) v'(R) \\ &- (p - \pi_{SI}) \rho v'(R + [(1 - \alpha)] \Pi) + (1 - p - q + \pi_{SI}) \rho v'(R + \Pi) < 0. \end{aligned}$$

En réarrangeant, on obtient :

$$\begin{aligned}
& - [(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) \\
& + (1 - p - q + \pi_{SI}) v' (R + \Pi) \\
& + (\pi_{SI} + (q - \pi_{SI})) v' (R)] \rho < - [\pi_{SI} + (q - \pi_{SI})] v' (R)
\end{aligned}$$

soit

$$\begin{aligned}
& [(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) \\
& (1 - p - q + \pi_{SI}) v' (R + \Pi) + qv' (R)] \rho > qv' (R)
\end{aligned}$$

et donc

$$\rho > \frac{qv' (R)}{(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) + (1 - p - q + \pi_{SI}) v' (R + \Pi) + qv' (R)} \quad (2.66)$$

On sait que $\rho \geq q$ pour avoir la rentabilité de l'assurance. Les situations sans espoir pour l'assureur sont donc telles que $\rho = q$ soit :

$$q > \frac{qv' (R)}{(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) + (1 - p - q + \pi_{SI}) v' (R + \Pi) + qv' (R)} \quad (2.67)$$

Par suite, en faisant $\rho = q$ dans (2.67), on a

$$1 > \frac{v' (R)}{(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) + (1 - p - q + \pi_{SI}) v' (R + \Pi) + qv' (R)}$$

et donc

$$(p - \pi_{SI}) v' (R + (1 - \alpha) \Pi) + (1 - p - q + \pi_{SI}) v' (R + \Pi) > (1 - q) v' (R)$$

ou encore

$$\left[\frac{v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} \right] > \left(\frac{1 - q}{p - \pi_{SI}} \right). \quad (2.68)$$

On peut remarquer que $0 < \forall \alpha < 1$,

$$\left[\frac{v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} \right] < 1$$

car $v'[R + (1 - \alpha)\Pi] > v'(R)$, $0 < \forall \alpha < 1$. De ce fait, la condition (2.68) est vérifiée que si en général, on a aussi

$$\frac{1 - q}{p - \pi_{SI}} < 1$$

ou encore

$$p + q > 1 + \pi_{SI} \quad (2.69)$$

Pour résumer, les situations où il est impossible d'avoir une assurance inondation rentable et où les exploitants ne font jamais l'effort d'arroser sont telles que :

$$\begin{cases} \frac{1 - q}{p - \pi_{SI}} < \left[\frac{v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} \right] < 1, & \text{(condition 1)} \\ \frac{\pi_{SI}}{p - \pi_{SI}} > \left[\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right] \left(\frac{\alpha\Pi - c}{c} \right). & \text{(condition 2)} \end{cases}$$

De la condition 2, on obtient

$$\pi_{SI} > \left[\frac{(\alpha\Pi - c)v'(R + (1 - \alpha)\Pi)}{cv'(R) + (\alpha\Pi - c)v'(R + (1 - \alpha)\Pi)} \right] p$$

ou de manière équivalente

$$\begin{aligned} p - \pi_{SI} &< p \left(\frac{cv'(R)}{cv'(R) + (\alpha\Pi - c)v'(R + (1 - \alpha)\Pi)} \right) \\ \frac{1 - q}{p - \pi_{SI}} &> \left(\frac{1 - q}{p} \right) \left(\frac{cv'(R) + (\alpha\Pi - c)v'(R + (1 - \alpha)\Pi)}{cv'(R)} \right) \\ &= \left(\frac{1 - q}{p} \right) \left[1 + \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \left(\frac{\alpha\Pi - c}{c} \right) \right] \end{aligned} \quad (2.70)$$

de sorte que, pour que les deux conditions soient vérifiées, il faut que :

$$\begin{aligned} \left(\frac{1-q}{p}\right) \left[1 + \left(\frac{v'(R + (1-\alpha)\Pi)}{v'(R)}\right) \left(\frac{\alpha\Pi - c}{c}\right)\right] &< \frac{1-q}{p - \pi_{SI}} \\ &< \frac{[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} < 1 \end{aligned} \quad (2.71)$$

sous la condition générale (2.69) liant p , q et π_{SI} .

On note en particulier, d'après la nouvelle expression de la condition 2 (condition 2.70), que cette condition est impossible si

$$\frac{1-q}{p} \geq 1$$

c'est-à-dire si $p + q \leq 1$. Par conséquent, pour que ces deux conditions soient vérifiées il faut donc que $1 - q < p$.

2.A.1.5 *Impossibilité d'avoir une assurance inondation rentable en situation de zéro effort d'arrosage*

D'après la condition 2.71, il ne serait pas possible d'avoir une assurance inondation rentable en situation de *zéro effort d'arrosage* si on a toujours

$$\begin{aligned} \left[1 + \left(\frac{v'(R + (1-\alpha)\Pi)}{v'(R)}\right) \left(\frac{\alpha\Pi - c}{c}\right)\right] \\ \times \left(\frac{1-q}{p}\right) < \frac{[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} \end{aligned}$$

ou encore si

$$\begin{aligned} \left(\frac{1-q}{p}\right) \left(\frac{\alpha\Pi - c}{c}\right) \left(\frac{v'(R + (1-\alpha)\Pi)}{v'(R)}\right) < \frac{[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} \\ - \frac{1-q}{p} \end{aligned}$$

a) Soit

$$\begin{aligned}
\mathcal{D}(\alpha, q) &= \left(\frac{1-q}{p} \right) \left(\frac{\alpha\Pi - c}{c} \right) \frac{v'(R + (1-\alpha)\Pi)}{v'(R)} \\
&\quad - \frac{[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} + \frac{1-q}{p} \\
&= \frac{(1-q)(\alpha\Pi - c)[v'(R) - v'(R + \Pi)]v'(R + (1-\alpha)\Pi)}{cpv'(R)[v'(R) - v'(R + \Pi)]} \\
&\quad - \frac{cpv'(R)[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{cpv'(R)[v'(R) - v'(R + \Pi)]} \\
&\quad - \frac{c(1-q)v'(R)[v'(R) - v'(R + \Pi)]}{cpv'(R)[v'(R) - v'(R + \Pi)]} \\
&= \frac{(1-q)[v'(R) - v'(R + \Pi)][(\alpha\Pi - c)v'(R + (1-\alpha)\Pi) + cv'(R)]}{cpv'(R)[v'(R) - v'(R + \Pi)]} \\
&\quad - \frac{cpv'(R)[v'(R + (1-\alpha)\Pi) - v'(R + \Pi)]}{cpv'(R)[v'(R) - v'(R + \Pi)]}.
\end{aligned}$$

b) On sait que ces conditions doivent être vérifiées $0 < \forall p, \forall q < 1 / (p + q) > 1$ et $\alpha > \frac{c}{\Pi}$.

Etudions la limite de $\mathcal{D}(\alpha, q)$ pour $(c/\Pi) < \forall \alpha < 1$ quand $q \rightarrow 1 - p, 1$:

$$\begin{aligned}
\lim_{q \rightarrow 1} \mathcal{D}(\alpha, q) &= - \frac{v'(R + (1-\alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)}. \\
&< 0.
\end{aligned}$$

Car, comme $v'(\cdot) > 0$, on n'en déduit

$$\frac{v'(R + (1-\alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} > 1.$$

Ceci vient du fait qu'étant donné que $v'(\cdot) > 0$, pour $(c/\Pi) < \forall \alpha < 1$,

$$v'[R + (1-\alpha)\Pi] > v'(R)$$

et donc

$$v'[R + (1-\alpha)\Pi] - v'(R + \Pi) > v'(R) - v'(R + \Pi).$$

Par suite

$$\frac{v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} > 1.$$

De même, pour $(c/\Pi) < \forall \alpha < 1$,

$$\begin{aligned} \lim_{q \rightarrow 1-p} \mathcal{D}(\alpha, q) &= \frac{cpv'(R) [v'(R) - v'(R + \Pi)]}{cpv'(R) [v'(R) - v'(R + \Pi)]} \\ &\quad - \frac{cpv'(R) [v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)]}{cpv'(R) [v'(R) - v'(R + \Pi)]}. \\ &= \frac{[v'(R) - v'(R + \Pi)] - [v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)]}{v'(R) - v'(R + \Pi)} \\ &= 1 - \frac{v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)}{v'(R) - v'(R + \Pi)} \\ &< 0 \end{aligned}$$

De plus, pour $0 < p < 1$ et $\alpha\Pi > c$

$$\frac{d\mathcal{D}(\alpha, q)}{dq} \Big|_{\forall q \in]1-p, 1]} = -\frac{1}{p} \left[1 + \left(\frac{\alpha\Pi - c}{c} \right) \frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right] < 0.$$

Par conséquent, pour $0 < \forall p, \forall q < 1 / (p + q) > 1$, il est possible d'avoir

$$\begin{aligned} \left[1 + \left(\frac{v'(R + (1 - \alpha)\Pi)}{v'(R)} \right) \left(\frac{\alpha\Pi - c}{c} \right) \right] \\ \times \left(\frac{1 - q}{p} \right) &< \frac{1 - q}{p - \pi_{SI}} \\ &< \frac{[v'(R + (1 - \alpha)\Pi) - v'(R + \Pi)]}{[v'(R) - v'(R + \Pi)]} \end{aligned}$$

tant que que $q \in]1 - p, 1]$.

2.A.2 Résultats des illustrations

2.A.2.1 La paire effort-assurance (θ^*, A^*) optimale et courbes de niveau (q, p) .

.

En utilisant l'équation (2.43) donnant le couple (θ, A) optimal et la fonction d'utilité définie en (2.44), on obtient :

$$\begin{aligned}
& (q - \pi_{SI}) (1 - \rho) (R - \rho A + A)^{-\eta} \\
& + \left(1 - \rho \left(\frac{\alpha \Pi}{\alpha \Pi - c}\right)\right) \pi_{SI} (R - \rho A - c\theta + A)^{-\eta} \\
& - (1 - p - q + \pi_{SI}) \rho (R - \rho A + \Pi)^{-\eta} = 0
\end{aligned}$$

ce qui donne

$$\begin{aligned}
\left(1 - \rho \left(\frac{\alpha \Pi}{\alpha \Pi - c}\right)\right) \pi_{SI} (R - \rho A - c\theta + A)^{-\eta} \\
= (1 - p - q + \pi_{SI}) \rho (R - \rho A + \Pi)^{-\eta} \\
- (q - \pi_{SI}) (1 - \rho) (R - \rho A + A)^{-\eta}.
\end{aligned}$$

On détermine l'expression des courbes de niveau (q, p) . En utilisant l'expression donnée à l'équation (2.43), on obtient :

$$\begin{aligned}
\left(1 - \rho \left(\frac{\alpha \Pi}{\alpha \Pi - c}\right)\right) \pi_{SI} (R - \rho A - c\theta + A)^{-\eta} &= (1 - p + \pi_{SI}) \rho (R - \rho A + \Pi)^{-\eta} \\
&- q [\rho (R - \rho A + \Pi)^{-\eta} \\
&+ (1 - \rho) (R - \rho A + A)^{-\eta}] \\
&+ \pi_{SI} (1 - \rho) (R - \rho A + A)^{-\eta}.
\end{aligned}$$

En notant

$$D_q = \rho (R - \rho A + \Pi)^{-\eta} + (1 - \rho) (R - \rho A + A)^{-\eta},$$

on obtient finalement

$$\begin{aligned}
q &= (1 - p) \left(\frac{\rho}{D_q}\right) (R - \rho A + \Pi)^{-\eta} + \left(\frac{\pi_{SI} \rho}{D_q}\right) (R - \rho A + \Pi)^{-\eta} \\
&- \left(\frac{(1 - \rho \left(\frac{\alpha \Pi}{\alpha \Pi - c}\right)) \pi_{SI}}{D_q}\right) (R - \rho A - c\theta + A)^{-\eta} \\
&+ \left(\frac{\pi_{SI} (1 - \rho)}{D_q}\right) (R - \rho A + A)^{-\eta} \tag{2.72}
\end{aligned}$$

2.A.2.2 Effort d'arrosage θ^* et niveau d'assurance A^* qui maximisent l'utilité espérée de l'exploitant

On considère les *CPOs* en θ et A données respectivement aux équations (2.33) et (2.34). Soit $H_{\theta,A}$ et $\delta_{\theta,A}$ respectivement la Matrice Hessienne et le Hessien tels que

$$H_{\theta,A} = \begin{pmatrix} \frac{\partial^2 E(V)}{\partial \theta^2} & \frac{\partial^2 E(V)}{\partial A \partial \theta} \\ \frac{\partial^2 E(V)}{\partial \theta \partial A} & \frac{\partial^2 E(V)}{\partial A^2} \end{pmatrix}$$

et

$$\delta_{\theta,A} \equiv |H_{\theta,A}| = \left(\frac{\partial^2 E(V)}{\partial \theta^2} \right) \left(\frac{\partial^2 E(V)}{\partial A^2} \right) - \left(\frac{\partial^2 E(V)}{\partial \theta \partial A} \right)^2.$$

La solution intérieure du système formé par les *CPOs* (2.33) et (2.34) est un maximum (relatif) si et seulement si $\delta_{\theta,A} > 0$ et $\frac{\partial^2 E(V)}{\partial \theta^2} < 0$.

Les dérivées secondes déterminées à partir des *CPOs* (2.33) et (2.34) s'écrivent :

$$\begin{aligned} \frac{\partial^2 E(V)}{\partial \theta^2} &\equiv \frac{\partial}{\partial \theta} \left(\frac{\partial E(V)}{\partial \theta} \right) \\ &= \pi_{SI} c^2 v''(R - \rho A - c\theta + A) \\ &\quad + (p - \pi_{SI}) [\alpha \Pi - c]^2 v''(R - \rho A + [\alpha \theta + (1 - \alpha)] \Pi - c\theta) \\ &< 0, \quad \text{car } v''(\cdot) < 0; \end{aligned}$$

de même,

$$\begin{aligned} \frac{\partial^2 E(V)}{\partial A^2} &\equiv \frac{\partial}{\partial A} \left(\frac{\partial E(V)}{\partial A} \right) \\ &= (1 - \rho)^2 [\pi_{SI} v''(R - \rho A - c\theta + A) + (q - \pi_{SI}) v''(R - \rho A + A)] \\ &\quad + \rho^2 (p - \pi_{SI}) v''(R - \rho A + [\alpha \theta + (1 - \alpha)] \Pi - c\theta) \\ &\quad + \rho^2 (1 - p - q + \pi_{SI}) v''(R - \rho A + \Pi) \end{aligned}$$

et

$$\begin{aligned} \frac{\partial^2 E(V)}{\partial A \partial \theta} &\equiv \frac{\partial}{\partial A} \left(\frac{\partial E(V)}{\partial \theta} \right) = \frac{\partial}{\partial \theta} \left(\frac{\partial E(V)}{\partial A} \right) \\ &= -c\pi_{SI} (1 - \rho) v'' (R - \rho A - c\theta + A) \\ &\quad - (p - \pi_{SI}) [\alpha \Pi - c] \rho v'' (R - \rho A + [\alpha \theta + (1 - \alpha)] \Pi - c\theta). \end{aligned}$$

Notons

$$\begin{aligned} v''_{SI} &= v'' (R - \rho A - c\theta + A) \\ v''_{\emptyset I} &= v'' (R - \rho A + A) \\ v''_{S\emptyset} &= v'' (R - \rho A + [\alpha \theta + (1 - \alpha)] \Pi - c\theta) \\ v''_{\emptyset\emptyset} &= v'' (R - \rho A + \Pi) \end{aligned}$$

Par suite,

$$\begin{aligned} \delta_{\theta,A} &= \{ (1 - \rho)^2 [\pi_{SI} v''_{SI} + (q - \pi_{SI}) v''_{\emptyset I}] \cdot \\ &\quad + \rho^2 [(p - \pi_{SI}) v''_{S\emptyset} + (1 - p - q + \pi_{SI}) v''_{\emptyset\emptyset}] \} \\ &\quad \times [\pi_{SI} c^2 v''_{SI} + (p - \pi_{SI}) (\alpha \Pi - c)^2 v''_{S\emptyset}] \\ &\quad - \{ [c\pi_{SI} (1 - \rho) v''_{SI} + (p - \pi_{SI}) [\alpha \Pi - c] \rho v''_{S\emptyset}]^2 \} \\ &= \{ [(1 - \rho)^2 \pi_{SI} v''_{SI} + \rho^2 (p - \pi_{SI}) v''_{S\emptyset}] \\ &\quad \times [\pi_{SI} c^2 v''_{SI} + (p - \pi_{SI}) (\alpha \Pi - c)^2 v''_{S\emptyset}] \} \\ &\quad + \{ [(1 - \rho)^2 (q - \pi_{SI}) v''_{\emptyset I} + \rho^2 (1 - p - q + \pi_{SI}) v''_{\emptyset\emptyset}] \\ &\quad \times [\pi_{SI} c^2 v''_{SI} + (p - \pi_{SI}) (\alpha \Pi - c)^2 v''_{S\emptyset}] \} \\ &\quad - \{ [c\pi_{SI} (1 - \rho) v''_{SI} + (p - \pi_{SI}) [\alpha \Pi - c] \rho v''_{S\emptyset}]^2 \} \end{aligned}$$

et donc

$$\begin{aligned}
\delta_{\theta,A} &= \left\{ [c(1-\rho)\pi_{SI}v''_{SI}]^2 + (1-\rho)^2\pi_{SI}(p-\pi_{SI})(\alpha\Pi-c)^2v''_{S\emptyset}v''_{SI} \right. \\
&\quad \left. + \rho^2(p-\pi_{SI})\pi_{SI}c^2v''_{S\emptyset}v''_{SI} + [\rho(p-\pi_{SI})(\alpha\Pi-c)v''_{S\emptyset}]^2 \right\} \\
&\quad - \left\{ [c(1-\rho)\pi_{SI}v''_{SI}]^2 + [\rho(p-\pi_{SI})(\alpha\Pi-c)v''_{S\emptyset}]^2 \right. \\
&\quad \left. + 2[c(1-\rho)\pi_{SI}\rho(p-\pi_{SI})(\alpha\Pi-c)v''_{S\emptyset}v''_{SI}] \right\} \\
&\quad + \left\{ [(1-\rho)^2(q-\pi_{SI})v''_{\emptyset I} + \rho^2(1-p-q+\pi_{SI})v''_{\emptyset\emptyset}] \right. \\
&\quad \left. \times [\pi_{SI}c^2v''_{SI} + (p-\pi_{SI})(\alpha\Pi-c)^2v''_{S\emptyset}] \right\} \\
&= \left\{ \pi_{SI}(p-\pi_{SI})v''_{S\emptyset}v''_{SI} [\rho^2c^2 + (1-\rho)^2(\alpha\Pi-c)^2 - 2c(1-\rho)\rho(\alpha\Pi-c)] \right\} \\
&\quad + \left\{ [(1-\rho)^2(q-\pi_{SI})v''_{\emptyset I} + \rho^2(1-p-q+\pi_{SI})v''_{\emptyset\emptyset}] \right. \\
&\quad \left. \times [\pi_{SI}c^2v''_{SI} + (p-\pi_{SI})(\alpha\Pi-c)^2v''_{S\emptyset}] \right\}.
\end{aligned}$$

Enfin,

$$\begin{aligned}
\delta_{\theta,A} &= \left\{ \pi_{SI}(p-\pi_{SI})v''_{S\emptyset}v''_{SI} [\rho c - (1-\rho)(\alpha\Pi-c)]^2 \right\} \\
&\quad + \left\{ [(1-\rho)^2(q-\pi_{SI})v''_{\emptyset I} + \rho^2(1-p-q+\pi_{SI})v''_{\emptyset\emptyset}] \right. \\
&\quad \left. \times [\pi_{SI}c^2v''_{SI} + (p-\pi_{SI})(\alpha\Pi-c)^2v''_{S\emptyset}] \right\} \\
&> 0, \quad \text{car } v''(\cdot) < 0.
\end{aligned}$$

Comme $\frac{\partial^2 E(V)}{\partial \theta^2} < 0$ et $\delta_{\theta,A} > 0$, par conséquent, le couple (θ^*, A^*) satisfaisant le système formé par les *CPOs* (2.33) et (2.34) est un maximum (relatif).

Essai # 3

Accès aux biens et services des ménages
aux revenus les plus faibles : un modèle
microéconomique de demande

Résumé. L'objectif initial de cet essai était d'étudier dans le cadre d'un modèle simple s'il existait des circonstances dans lesquelles une entreprise productrice de biens et services trouverait profitable de ne pas exclure de consommateurs. La problématique générale consistait à essayer de faire le lien entre distribution des revenus au sein de la population, niveau des prix de marché et accès aux biens et services pour les ménages les plus pauvres.

Pour ce faire, le cadre d'analyse est d'abord restreint au cas dit de « la demande unitaire » c'est-à-dire au cas où les individus soit se portent acquéreurs d'une seule unité du bien, soit préfèrent s'en dispenser entièrement. Afin de ré-introduire les effets revenus absents du modèle quasi-linéaire, nous étendons le modèle de demande unitaire classique. Pour ne pas introduire de biais dans l'analyse, nous continuons de supposer, à l'instar du modèle classique, que l'utilité dérivée de la possession du bien étudié est additivement séparable de celle dérivée de l'utilisation des autres biens. Ce faisant, le bien est aussi utile aux riches qu'aux pauvres, si bien que les différences de comportement ne sont pas imputables « aux préférences » mais à la seule « capacité à payer »

Dans ce cadre étendu, il est facile de montrer que, pour tout prix p , il existe un revenu $R_0(p)$ en deçà duquel les ménages vont trouver le bien trop onéreux. La demande agrégée n'est autre que la taille la population ayant un revenu au moins égal à $R_0(p)$. Les propriétés de la demande sont donc directement reliées à la distribution des revenus dans la population. Toute augmentation de prix se traduit par une contraction de la demande, toute diminution du prix par son expansion. Il est possible d'identifier des circonstances dans lesquelles la distribution de revenus est suffisamment concentrée vers les bas revenus pour que, même une entreprise qui cherche à maximiser ses profits ait intérêt à ce que ses produits et services soient « accessibles » à toute la population, (*i.e.* aient un prix suffisamment bas pour être acquis par tous).

A partir de la famille des distributions de Pareto, nous étudions ensuite le lien entre inégalités de revenu et « niveau des prix », dans un modèle où le prix d'équilibre est fixé par une entreprise en situation de monopole, qui maximise ses profits. Nous montrons que, dans le cadre de notre modèle, la marge du monopole diminue strictement avec le paramètre de Pareto et qu'elle tend vers zéro lorsqu'on s'approche de l'égalité parfaite des

revenus. En d'autres termes, les inégalités de revenu augmentent la marge réalisée par le monopole sur le marché - et, dans ce modèle de demande unitaire, sont à la source des inefficacités de la concurrence imparfaite.

Le modèle de demande unitaire est ensuite étendu pour autoriser deux niveaux de qualité du même bien - et le cas où ces deux qualités sont proposées par deux entreprises en concurrence est également considéré. Nous montrons, là-encore, que les prix d'équilibre augmentent avec le niveau d'inégalités et que ces dernières ne font pas que réduire le bien-être des consommateurs (diminuer leur surplus) mais également l'efficacité des marchés (elles augmentent la charge morte).

Pour examiner dans quelle mesure les résultats précédents découlent des spécificités du modèle de demande unitaire ou peuvent être considérés comme plus généraux, nous proposons ensuite un résultat technique permettant d'établir la Schur-convexité d'une fonction implicite. La Schur-convexité (ou concavité) d'une fonction est une propriété qui établit un lien entre sa monotonie et le degré d'inégalité de la distribution à laquelle elle s'applique. En d'autres termes, une fonction Schur-convexe (concave) diminue (augmente) avec les inégalités. La condition du premier ordre (CPO) dérivée de la maximisation des profits définit implicitement le prix de monopole. Nous appliquons notre lemme pour montrer que, dans l'exemple que nous considérons, ce prix d'équilibre est Schur-concave. Autrement dit, nous exhibons un modèle où les individus ont les mêmes préférences (qui donnent lieu à une fonction de demande individuelle « standard », décroissante en prix et croissante en revenu) et ne diffèrent que par leur revenus ; et nous montrons que la marge relative réalisée à l'équilibre par un monopole qui maximise ses profits constitue un index de mesure des inégalités de revenus de cette population.

Il ne s'agit pas évidemment d'une propriété universelle des prix de marchés, ne serait-ce parce qu'il existe des biens de luxe ou des biens de Giffen pour lesquels la demande fait montre de comportements opposés à ceux que nous avons supposés ici. Notre modèle établit cependant qu'il existe des circonstances assez larges (que notre lemme permet d'identifier) où les inégalités sont source d'inefficacités de marché - au sens où leur accroissement se traduit systématiquement par un pouvoir de marché accru des entreprises, et

donc des marges ainsi qu'une « charge morte » plus élevées. Dans ce cadre où les effets revenus ont été ré-introduits, toute redistribution qui amène une baisse des inégalités est donc génératrice d'efficacité - indépendamment de toute considération normative. Notre travail identifie une conséquence des inégalités de revenu jusqu'ici ignorée (à notre connaissance) et souligne les limites du cadre d'analyse quasi-linéaire pour tout ce qui a trait à leur étude.

3.1 Introduction

Il est largement admis et documenté que les conditions de vie des ménages à bas revenus s'expliquent en majeure partie par leurs faibles dotations initiales et les faibles ressources dont ils disposent. Dans un rapport assez densément documenté publié en 2010 pour le compte de "Office of Fair Trading (OFT)", "Europe Economics and New Policy Institute (EENPI)" a noté que les individus à faibles revenus ont moins de choix que les individus à revenus élevés. L'EENPI (2010) a précisé que non seulement, une proportion importante des ressources de ces consommateurs est consacrée à l'acquisition de biens essentiels à leur survie mais encore ils n'ont pas accès à un certain nombre de services clé capables de les aider à améliorer leur accès à d'autres marchés et augmenter leurs possibilités de choix. Du point de vue théorique, Bradshaw (2007) a synthétisé les principales causes de la pauvreté individuelle (dotations initiales, incompetence, ...), culturelle (valeurs contraires au progrès, ...), politico-économique (structures économiques et politiques créant des barrières à l'accès pour les pauvres,...), géographique (des zones géographiquement défavorisées) cumulative et cyclique (spirale de pauvreté,...). Blank (2003) a classé et synthétisé en cinq approches les cadres théoriques d'explication de la pauvreté. Selon ces cinq approches, la pauvreté est liée à l'absence de marchés fonctionnant efficacement, à certaines caractéristiques empêchant certains individus de participer de manière productive au fonctionnement du marché. A cela s'ajoutent les défaillances du marché (asymétries d'information, externalités, ...), des processus sociaux et politiques (corruption, instabilité chronique, ...) exogènes au marché et qui impactent celui-ci et enfin les caractéristiques comportementales des individus (passivité, ...) et de leurs choix (famille nombreuse, ...). Dans un article publié en 2013 par Son (2013), la pauvreté des ménages constitue l'un des facteurs importants de nonaccès à l'éducation de base et aux services infra-structurels dans les pays en développement de l'Asie.

Un bon nombre de recherches ont permis d'identifier les différents facteurs responsables et les diverses manifestations de cet accès très limité de ces individus aux biens et services (Deutsch *et al.*, 2012; EENPI, 2010). D'autres ont établi et caractérisé les conditions dans lesquelles un marché, faisant face à une demande hétérogène liée à une

différence significative de revenus, peut être approvisionné (He et Sun, 2006; Malueg et Schwartz, 1994). Vu l'importance des revenus dans ce qu'il en est des conditions de vie des gens, il s'avère donc que la problématique de l'accès aux biens et services des individus à faibles revenus doit en partie passer, étant donné leurs préférences, par l'analyse de leur structure de revenus et la distribution des revenus des individus. Dans l'optique d'une petite contribution tout au moins théorique à l'amélioration de l'accès des individus à faibles revenus aux biens et services, nous nous proposons de répondre à cette :

Question. *La recherche des profits par une entreprise productrice de biens et services mène-t-elle toujours à l'exclusion ?*

Nous répondrons à cette question en élaborant un modèle microéconomique devant nous aider dans un premier temps à identifier les circonstances sous lesquelles, la recherche du profit conduit les entreprises à adopter naturellement une politique tarifaire qui donne accès aux biens à tous.

De ces considérations, la démarche adoptée pour développer le modèle est la suivante. Nous commençons d'abord par présenter les principaux résultats du modèle de manière à bien structurer son développement. Dans un second temps, nous développons le modèle. Ce développement s'organise de la façon suivante. Les préférences des consommateurs, supposées par la suite non quasi-linéaires, sont définies et modélisées. De la définition de ces préférences, la demande pour un nombre n (R) de consommateurs ayant un revenu R est déterminée ; les profits de l'entreprise productrice et le bien-être social calculés pour cette demande sont ensuite évalués. S'ensuivront, sous certaines hypothèses simplificatrices, une analyse de comment le prix du bien est fixé sous l'angle de la maximisation des profits et du bien-être social et une évaluation des conséquences sur le niveau d'exclusion. Les résultats établis sont illustrés en utilisant tour à tour la distribution de Pareto de type 1 et la distribution log-normale. Dans un troisième temps, les résultats du modèle sont généralisés et une relation liant les préférences des ménages à la distribution des revenus est établie. Ces résultats sont confrontés à la littérature. Le cadre général de référence du modèle est précisé. Enfin, on terminera ce chapitre par une synthèse des principaux résultats trouvés, les leçons qu'on peut en tirer, les limites du modèle et les perspectives.

Encadré 3.1 : Haïti : Accès et Exclusion

Les données d'enquête disponibles relatives à la pauvreté et à l'accès aux biens et services par les haïtiens datent de l'Enquête, réalisée en 2012, sur les Conditions de Vie des Ménages Après le Séisme (ECVMAS) de 2010. La réalisation d'enquête de ce genre n'a pu avoir lieu à cause de l'instabilité socio-politique chronique qu'a connue le pays depuis et qui s'est vue exacerbée depuis la prolongation contestée du mandat du président Jovenel Moïse culminant avec son assassinat dans la nuit du 7 juillet 2021.

D'après les données de l'ECVMAS, près d'un quart (24%) de la population haïtienne (tableau suivant) vit dans l'extrême pauvreté dont 38% en milieu rural, le taux de pauvreté s'élevant à 59%. L'indice de Gini de 0.61 traduit une très forte inégalité de revenus dans le pays. L'instabilité socio-politique aigüe et l'insécurité incomparable qu'a connues le pays contribuent vraisemblablement à aggraver la situation.

Haïti: Taux de Pauvreté

	Pauvreté Extrême (Ligne: 1\$US)	Pauvreté (Ligne: 2\$US)	Coeff. Gini	Pauvreté Extrême		
				Rurale	Aire Mét.	Autre Urb.
EBCM 2000	31%	<i>na</i>	<i>na</i>	38%	20%	21%
ECVMAS 2012	24%	59%	0.61	38%	5%	12%

Source: Institut Haïtien de Statistique et d'Informatique (IHSI)

EBCM: Enquête Budget-Consommation des Ménages

ECVMAS: Enquête sur les Conditions de Vie des Ménages Après le Séisme de 2010

Note:

Les 20 % les plus riches de sa population détiennent plus de 64 % de sa richesse totale, tandis que les 20 % les plus pauvres en possèdent moins de 1 % (Banque Mondiale, juin 2022).

En termes d'accès aux biens et services, les rares données disponibles (tableau en dessous) ont permis de constater que 66.7% possèdent uniquement un réchaud à charbon de bois. Respectivement 20% et 50.7% des ménages font leur besoin et évacuent leurs déchets solides dans « la nature ». Pour s'éclairer, 53.5% des ménages ont recours à des lampes à kérozène.

Haïti: Accès aux biens et services**Possession (% ménages) de certains biens ou équipements fonctionnels**

Réchaud à charbon de bois	66.7%
Four	5.5%
Ordinateur	3.9%
Connexion internet	1.6%
Autres	22.3%

Possession de lieu d'aisance (toilette) (% ménages)

Latrine non améliorée	27.4%
Latrine améliorée	22.9%
"La Nature"	20.0%
Autres	29.7%

Mode d'évacuation des déchets solides (% ménages)

SMCRS	9.7%
"La Nature"	50.7%
Autres	39.6%

Source d'énergie (éclairage) (% ménages)

Electricité d'Haïti (EDH)	11.5%
Lampe à gaz	53.5%
Autres	35.0%

Source: IHSI, ECVMAS 2012

SMCRS: Service Métropolitain de Collecte des Résidus Solides

Mode d'évacuation des eaux usées

Système d'égouts	19.8%
"La Nature"	44.5%
Autres	35.5%

Encadré 3.1 : Haïti : Accès et Exclusion (suite)

D'après les résultats de l'enquête FinScope réalisée en 2018 par la Banque de la République d'Haïti (BRH), si l'on considère le nombre de personnes qui utilisent les services des institutions financières (bancaires ou non bancaires), 46% de la population haïtienne adulte est financièrement exclue. Il existe également des disparités selon le milieu de résidence. Dans le milieu urbain, 12% de la population est bancarisée contre 4% en milieu rural. Un pourcentage de 36% de la population utilisent les services des institutions financières autres que les banques contre 21% en milieu rural. Les services financiers informels sont utilisés par 8% de la population en milieu urbain contre 18% en milieu rural ; et en milieu urbain, 44% de la population est exclue financièrement contre 57% en milieu rural. Quant à l'accès au crédit, 4% de la population adulte haïtienne ont recours au crédit formel (banques et autres institutions financières), 5% ont recours au crédit informel, 5% empruntent auprès de la famille/ami et 87% n'empruntent pas.

3.2 Les principaux résultats du modèle

Le but principal de cet essai, comme précisé précédemment, est d'identifier les circonstances sous lesquelles, la recherche du profit conduit les entreprises à adopter naturellement une politique tarifaire qui donne accès aux biens à tous. Nous établissons en effet que, pour la famille de distributions de revenus que nous considérons, lorsque les écarts de revenus au sein de la population sont suffisamment faibles, même une entreprise qui ne cherche qu'à faire des profits va adopter une politique tarifaire qui rend ses biens ou services accessibles à tous. Ce résultat est maintenu si on suppose que deux qualités - produites par un monopole ou un duopole formé de deux firmes compétitives - du bien sont disponibles.

La proposition 3.11 formulée dans le cadre du développement de notre modèle, établit sans équivoque que le prix optimal du bien p^* maximisant le profit s'ajuste en fonction de l'élasticité-prix de la demande des ménages ayant le revenu minimum \underline{R} de la société. De manière précise, l'écart relatif de prix s'ajuste en fonction du degré de sensibilité des consommateurs au prix. Dans un second temps, on a établi l'existence d'un revenu minimal R_0 donnant accès au bien (proposition 3.7). Ce revenu minimal est une fonction croissante du prix p (proposition 3.8). L'augmentation induite de R_0 est plus que proportionnelle à l'accroissement de p du fait que le ménage valorise davantage le bien en ce sens qu'il retire une utilité plus importante, proportionnelle à l'augmentation du revenu,

à l'achat d'une unité du bien et augmente plus que proportionnellement à l'accroissement du prix p . Nous avons pu établir aussi la Schur-convexité (proposition 3.9) de la fonction obtenue des conditions du premier ordre de la fonction de profits. Cette propriété garantit qu'au prix p fixé, la fonction Schur-convexe croît avec les inégalités. Enfin, nous avons pu déterminer que prix unique p vérifiant $R_0(p) = \underline{R}$ est déterminé dans le cadre d'une distribution des revenus telle que le monopole trouve profitable de ne pas exclure de consommateurs (proposition 3.10). Ce résultat prouve bien qu'en partie les caractéristiques de la distribution des revenus déterminent la façon dont les prix se forment dans l'économie. Et donc, une modification rendant la distribution des revenus plus égalitaire peut se révéler efficace pour donner accès au bien à tous ceux ayant un revenu minimum \underline{R} et qui le désirent.

Naturellement, si une entreprise fait face à une population où la distribution des revenus est très inégalitaire, il lui sera au contraire profitable d'exclure certains consommateurs potentiels pour pouvoir s'assurer une marge plus importante sur les autres. Les illustrations effectuées, en supposant que les revenus sont Pareto-distribués ou log normale-distribués, ont permis d'étayer l'idée que d'autant plus inégalitaire est la répartition des revenus et d'autant plus élevé est le prix optimal p^* du bien résultant de la maximisation du profit.

En d'autres termes, notre modèle micro-économique pointe sur un bénéfice supplémentaire des politiques de redistribution. Au delà de leur impact direct sur les ménages, elles peuvent en effet ré-orienter la dynamique de création de richesses associée aux économies de marché pour les rendre inclusives - et non pas sources de davantage d'inégalités.

3.3 Le Cadre général

Ce modèle de demande s'inscrit dans un cadre de référence utilitariste classique. Le choix de ce cadre résulte de ce que l'utilitarisme constitue une famille de philosophies morale et politique selon lesquelles le bien-être social est en fin de compte la seule valeur éthique ou le seul bien à maximiser. L'utilitarisme, basé sur le principe d'utilité visant « le plus grand bonheur du plus grand nombre », prescrit, dans sa version la plus utilisée par

les économistes, que le bien-être ¹ d'une société est maximisé pour un résultat ("outcome") donné quand celui-ci maximise la somme des utilités des individus de cette société (Riley, 2016; Baujard, 2013a; Shafer-Landau, 2016). Le terme "utile" désigne alors ce qui contribue à maximiser le bien-être d'une communauté (Bozzo-Rey et Dardenne, 2012). Dans ce cadre là, la distribution des niveaux d'utilité (revenu) est sans impact sur le bien-être social. En effet, la fonction de bien-être social utilitariste mesurant le bien-être social comme étant la somme des utilités (revenus) individuelles, est maximale quand l'est la somme des utilités au niveau de la société indépendamment de la distribution des revenus au niveau de cette société (Baujard, 2013b). En conséquence, la présence – ou non – d'inégalités n'appelle *a priori pas* à une quelconque intervention.

3.4 Description du modèle

Dans cette section, nous élaborons le modèle dans lequel nous identifierons d'abord les circonstances où la recherche du profit amène spontanément les entreprises à adopter une politique tarifaire inclusive. Dans cette démarche, la problématique de l'accès aux biens et services est traitée, non pas comme un problème émanant 'de la demande', mais comme résultant possiblement du comportement 'de l'offre'. Ensuite, nous établissons le cadre d'une distribution des revenus dans lequel le monopole trouve profitable de ne pas exclure de consommateurs.

3.4.1 Préférences des ménages et principales hypothèses du modèle

On introduit une représentation stylisée de l'économie en supposant qu'il n'y a que deux « biens ». Il y a d'une part le bien que l'on étudie, X , et d'autre part, un bien composite, le revenu disponible R pour « toutes les autres dépenses ». Ce bien composite constitue une variable importante à prendre en compte pour modéliser les préférences des

1. Maskin (1978) a construit, en considérant une fonction de bien-être social à la d'Aspremont et Gevers (1977) et en y introduisant l'axiome de comparabilité complète, un théorème dans lequel il a établi le principe fondamental de l'Utilitarisme.

ménages à faibles revenus.

Cette représentation stylisée de l'économie nous permettra de « séparer » le bien X pour lequel nous nous proposerons de déterminer la demande, du bien composite R représentant « toutes les autres dépenses ». Évidemment, cette démarche d'agrégation tout en aidant à isoler le bien que l'on veut étudier, constitue une abstraction importante simplifiant la réalité.

Les préférences des ménages . Les caractéristiques individuelles observables des consommateurs, en particulier leur niveau de revenu, fournissent des indications importantes sur le type de préférences à utiliser pour modéliser leur comportement. De ce fait, pour modéliser les préférences des individus à faibles revenus, nous adoptons une *fonction d'utilité qui n'est pas quasi-linéaire* $\mathcal{U}(x, R)^2$ supposée séparable³, croissante, dérivable à l'ordre deux et concave dans chacun de ses deux arguments. Formellement, on a :

$$\frac{\partial \mathcal{U}}{\partial x} \geq 0, \quad \frac{\partial^2 \mathcal{U}}{\partial x^2} \leq 0, \quad (3.1)$$

$$\frac{\partial \mathcal{U}}{\partial R} > 0, \quad \frac{\partial^2 \mathcal{U}}{\partial R^2} < 0. \quad (3.2)$$

Les inégalités (3.1) ne font que formaliser les hypothèses standards faites quant à l'utilité marginale d'un bien qui est donc supposée (faiblement) décroissante. Cela formalise l'idée qu'un bien est d'autant moins utile qu'on en possède déjà une grande quantité. Les inégalités (3.2), bien que procédant de la même logique, nous mettent en dehors du cadre

2. Ce choix sous-tend que les inégalités d'accès aux biens et services induites par la faiblesse du revenu s'avèrent inefficaces et ont des conséquences négatives sur le niveau de bien-être (Schiff, 2004). L'adoption d'un tel cadre n'implique pas à priori l'hypothèse d'existence de préférences sociales pour la redistribution.

3. La notion de préférences séparables est très souvent utilisée par les économistes quand ils étudient un bien ou un groupe de biens spécifiques. L'idée est de « séparer » ce bien ou ce groupe de biens de tous les « autres biens » représentant un bien composite dans un modèle à « deux biens ».

quasi-linéaire qui postule l'absence de tout effet revenu.⁴ Ces inégalités impliquent aussi qu'une augmentation marginale du revenu procure davantage de bien-être à un individu à revenus faibles qu'à un individu avec des revenus plus importants⁵.

L'utilité d'un consommateur de revenu R qui acquiert la quantité x du bien X en échange d'un « transfert monétaire » $t(x)$, s'écrit donc

$$U(x, R) = u[x; R - t(x)],$$

où l'on suppose, par simplicité que $t(x)$ est une fonction continue, croissante et dérivable.

Les principales hypothèses. Afin de mieux cerner la problématique qui nous intéresse, nous ajoutons plusieurs hypothèses simplificatrices à la représentation stylisée introduite jusqu'ici.

Hypothèse 3.1. *Le bien X n'est pas un bien essentiel*

On suppose tout d'abord que le bien X n'est pas un bien essentiel⁶. En d'autres termes, on ne fait pas l'hypothèse d'Inada (Inada, 1963) pour supposer au contraire que

$$\lim_{X \rightarrow 0^+} \frac{\partial U}{\partial X} = \bar{u} < +\infty.$$

4. Pour leur simplicité, les préférences quasi-linéaires sont surtout utilisées en économie du bien-être sous l'hypothèse implicite que les niveaux de revenus sont assez élevés. Sous cette hypothèse, une augmentation de revenus a très peu d'effet sur la demande (Varian, 1984). Par contre, dans le cas où les individus concernés ont de faibles niveaux de revenus, une augmentation des revenus a un impact non négligeable sur la demande. Clark (1973) était arrivé à cette conclusion :

« *Not only the marginal utility of \$1, but the marginal utility of one hour's earnings, may be expected to continue to fall with rising real income-if not, we should expect to find men working longer hours as their incomes rose. It is, however, probable that the rate of fall of marginal utility of one hour's earnings-and hence in the number of hours worked-will become very small* ».

5. Cette théorie est formalisée par Tirole (1988) au niveau du chapitre 2 de son ouvrage

6. Un bien essentiel est un bien pour lequel il existe une demande plutôt inélastique. La consommation d'un tel bien en quantité de plus en plus petites procure un bien-être de plus en plus intense. La définition de la notion de « biens essentiels » doit prendre en compte les caractéristiques géographiques, le niveau de développement et les spécificités historique et culturelle du pays en question (French-Davis, 1976).

Hypothèse 3.2. *Le bien R est un bien essentiel*

Par contre, nous supposons qu'il existe des biens essentiels à la survie des ménages. En conséquence, nous maintenons l'hypothèse d'Inada pour le revenu disponible R représentant « toutes les autres dépenses ». Le bien R peut donc être qualifié d'essentiel et donc :

$$\lim_{R \rightarrow 0^+} \frac{\partial \mathcal{U}}{\partial R} = +\infty.$$

Hypothèse 3.3. *L'utilité marginale décroît pour des revenus de plus en plus élevés*

On suppose également, pour des niveaux de revenu de plus en plus élevés, que l'utilité additionnelle retirée d'un accroissement infinitésimal du revenu R , est asymptotiquement nulle c'est-à-dire

$$\lim_{R \rightarrow +\infty} \frac{\partial \mathcal{U}}{\partial R} = 0.$$

Ces deux dernières hypothèses sur le comportement de $(\partial \mathcal{U} / \partial R)$ – ainsi que celle faite sur la « non-essentialité » de X – signifient qu'en dessous d'un certain seuil de revenu, les consommateurs préfèrent renoncer totalement au bien X mais qu'au contraire, il est toujours intéressant d'en acquérir s'ils disposent d'un revenu suffisant.

Hypothèse 3.4. *Le bien X n'est pas plus utile pour les riches que pour les pauvres*

Afin que le cadre d'analyse reste neutre quant à ses implications en termes (re-)distributifs⁷, on suppose également que

$$\frac{\partial^2 \mathcal{U}}{\partial R \partial X} = \frac{\partial}{\partial X} \left(\frac{\partial \mathcal{U}}{\partial R} \right) \equiv 0.$$

En d'autres termes, le bien X , « en soi », n'est ni plus utile ni moins utile pour les riches ou pour les pauvres.

7. Cette hypothèse de neutralité garantit l'absence sur les résultats du modèle de l'impact d'un certain nombre de variables influant sur les préférences des individus pour la redistribution. Une synthèse des principaux déterminants théoriques des préférences pour la redistribution et une évaluation empirique peuvent être trouvées dans un excellent chapitre d'ouvrage rédigé par Alesina et Giuliano (2011).

3.4.2 Utilité et demande

Dans ce qui suit, pour mieux isoler l'importance du revenu et de sa distribution dans l'accès au bien, nous restreignons le cadre d'analyse pour nous concentrer sur le cas dit de « la demande unitaire ». En d'autres termes, nous supposons que, sauf à y renoncer totalement, *il n'est pas possible d'acquérir moins qu'une unité du bien X* ; de plus, l'acquisition d'une quantité supplémentaire de ce même bien ne donne lieu à aucune augmentation de bien-être. En d'autres termes, nous nous plaçons dans un cadre où les individus soit se porteront acquéreurs d'une seule unité du bien, soit préféreront s'en dispenser entièrement.

3.4.2.1 Utilité et demande individuelle

En référence au type de préférences retenues (Sous-section 3.4.1) pour les ménages à faibles revenus, nous adoptons une fonction d'utilité séparable similaire à celle utilisée par He et Sun (2006) :

$$\mathcal{U}(x, R) = \ln(1 + u.I_{x \geq 1}) + \ln(R - p), \quad (3.3)$$

où : $u > 0$, p dénote le prix unitaire du bien X et $I_{x \geq 1}$ désignant la fonction indicatrice qui vaut 1 quand $x \geq 1$ et zéro sinon. Par définition, une personne préfère acheter une unité du bien X plutôt que de rester sans rien quand $\mathcal{U}(x = 1, R) > \mathcal{U}(0, R)$ (Gabszewicz et Thisse, 1979). Il s'ensuit qu'un individu se porte acquéreur du bien X si et seulement si :

$$\ln(1 + u) + \ln(R - p) > \ln(R).$$

De cette inégalité, en supposant R donné, le *prix unitaire maximal* p_0 auquel le bien X peut être acquis est tel que

$$p < \frac{u}{1 + u}R \equiv p_0(R). \quad (3.4)$$

Donc, pour R donné, un individu se porte acquéreur pour une unité du bien X tant que $p < p_0$. D'autant plus élevé est le revenu R , d'autant plus élevé est le prix unitaire maximal p_0 . De manière analogue, si on suppose p donné, le *revenu minimal* R_0 qui donne

lieu à l'achat du bien X est défini par :

$$R > p \left(1 + \frac{1}{u} \right) \equiv R_0(p). \quad (3.5)$$

Le revenu minimum R_0 tel que défini est *une fonction croissante* $R_0(p)$ de p . Un individu ayant un revenu R achète le bien X tant que $R > R_0$. De plus, un effort plus important doit être consenti en termes de revenu pour acquérir le bien X d'autant plus élevé est le prix unitaire.

3.4.2.2 Demande agrégée

On considère un monopoliste produisant un bien X et faisant face à un *nombre important* $n(R)$ de ménages ayant un revenu R . La fonction de distribution du nombre de ménages ayant un revenu R est formellement définie par : $n(R) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ⁸. Étant donné le revenu minimal R_0 donnant accès au bien X , le nombre d'unités du bien X achetées, *la demande agrégée* $D(p)$, est donné par :

$$D(p) \equiv D = \int_{R_0}^{+\infty} n(R) dR = N - N(R_0), \quad (3.6)$$

où $N(R)$ est la distribution cumulée associée à la fonction de densité $n(R)$ et N le nombre total de ménages tel que $N = \int_0^{+\infty} n(R) dR$.

On suppose, pour des raisons de simplicité, que les coûts totaux $C(D)$ de production auxquels fait face le monopoleur pour satisfaire la demande sont linéaires et définis par :

$$C(D) = cD. \quad (3.7)$$

En d'autres termes, il n'y a pas de coûts fixes et le coût unitaire de production du bien

8. Cette fonction est définie dans \mathbb{R}_+ en faisant l'hypothèse de l'existence d'un *continuum* (nombre très élevé) de consommateurs. Bien que critiquée pour son caractère irréaliste (Jablecki, 2007), cette hypothèse est adoptée dans divers travaux de recherche tels par exemple les articles de Braouezec (2012) et Malueg et Schwartz (1994), à des fins de dénombrement de population d'agents économiques de taille élevée.

X est noté $c > 0$.

3.4.3 Profits et Bien-être social

Les profits. Les profits (économiques) sont définis comme la différence entre les revenus perçus pour la vente d'un bien X produit et les coûts encourus (*Profits : encadré 3.4.3*).

Les profits $\Pi(p)$ du monopole produisant pour les $n(R)$ ménages qui achètent le bien sont alors définis par :

$$\begin{aligned}\Pi(p) &= (p - c) D(p) \\ &= \int_{R_0}^{+\infty} (p - c) n(R) dR.\end{aligned}\tag{3.8}$$

Encadré 3.4.3 : Profits.

Une entreprise, pour continuer à produire de manière viable, doit être capable de récupérer tout au moins les coûts de production encourus. Cette viabilité de la production de l'entreprise résulte de son efficacité économique c'est-à-dire de sa capacité à réaliser des résultats, des gains de productivité (efficacité économique) et d'améliorer sa rentabilité financière. L'importance des profits réalisés constitue donc un critère d'évaluation de cette efficacité économique ^a. Les profits (économiques) ^b sont définis comme la différence entre les revenus perçus pour la vente d'un bien X produit et les coûts encourus, les inputs utilisés et outputs obtenus étant évalués à leur coût d'opportunité. Si pour produire un bien X en quantité x , l'entreprise met en œuvre un ensemble d'actions qui lui coûtent $C(x)$ et qui lui procurent des revenus $I(x)$, ses profits s'élèvent alors à $\pi(x) = I(x) - C(x)$. L'entreprise qui agit toujours de manière à maximiser ses profits, continuera à produire le bien X de manière viable tant que $\pi(x) \geq 0$.

^a. En raison du risque et de l'incertitude liés à tout investissement donné, le profit, élément essentiel du processus d'accumulation du capital, est aussi défini comme la différence entre la rémunération ex-ante (anticipée dans un environnement de risques potentiels) et la rémunération ex-post (réalisée étant donné les risques effectifs encourus) du capital investi (Weston, 1954, 1950).

^b. Il faut ici différencier les profits économiques des profits comptables qui correspondent à la différence entre les revenus perçus par l'entreprise et les coûts monétaires encourus pour la production d'un bien X .

Le Bien-être social. Sous un angle utilitariste, le bien-être social est défini comme la somme des bien-être individuels ou alternativement comme la somme des utilités individuelles (Yew-Kwang, 1983). De ce qui précède, le bien-être social $W(p)$ est défini comme la somme des surplus des consommateurs W_C - qui achètent et qui n'achètent pas le bien - et du surplus (les profits) du producteur W_P .

Encadré 3.4.3 : Bien-être social.

Le bien-être social tel que défini et caractérisé par Yew-Kwang (1983) est un vecteur de bien-être individuels. Ici le bien-être individuel est pris dans le sens du bien-être subjectif ou bonheur subsumant simultanément plaisir sensuel et peine, délices et souffrances spirituelles. Alternativement, le bien-être social est défini comme un vecteur (de fonctions) des utilités individuelles ordinales représentant les préférences individuelles. Si de plus, au niveau du vecteur de bien-être individuels, le bien-être de quelques individus s'améliore sans que le bien-être de quelqu'un d'autre se détériore, le bien-être social s'améliore. Le bien-être social est alors défini au sens de Pareto. Sous un angle utilitariste, le bien-être social est défini comme la somme des bien-être individuels ou alternativement comme la somme des utilités individuelles (Yew-Kwang, 1983). C'est donc le total des richesses créées dans la société suite à la fabrication et à la commercialisation d'un bien donné.

Les consommateurs qui n'achètent pas le bien sont ceux ayant un revenu $R < R_0$.

Leur surplus W_C^1 est défini par :

$$W_C^1(p) = \int_0^{R_0} \ln(R) n(R) dR.$$

Le surplus W_C^2 des consommateurs qui achètent le bien - ceux ayant un revenu $R \geq R_0$ - est donné par :

$$W_C^2(p) = \int_{R_0}^{+\infty} [\ln(1+u) + \ln(R-p)] n(R) dR.$$

Le surplus total des consommateurs est donc donné par :

$$\begin{aligned} W_C(p) &= W_C^1(p) + W_C^2(p) \\ &= \int_0^{R_0} \ln(R) n(R) dR + \int_{R_0}^{+\infty} [\ln(1+u) + \ln(R-p)] n(R) dR. \end{aligned}$$

Du côté du producteur, le surplus (les profits) W_P est donné par :

$$W_P(p) = \int_{R_0}^{+\infty} (p-c) n(R) dR.$$

Le bien-être social ou surplus collectif $W(p)$ correspond donc à :

$$\begin{aligned}
 W(p) &= W_C(p) + W_P(p) \\
 &= \left\{ \int_0^{R_0} \ln(R) n(R) dR + \int_{R_0}^{+\infty} [\ln(1+u) + \ln(R-p)] n(R) dR \right\} \\
 &+ \left\{ \int_{R_0}^{+\infty} (p-c) n(R) dR \right\}. \tag{3.9}
 \end{aligned}$$

Dans ce cadre, on se propose de déterminer les circonstances telles que la recherche du profit conduit les entreprises à adopter naturellement une politique tarifaire qui donne accès aux biens à tous. Cette démarche se déroule sous un certain nombre de contraintes de faisabilité que nous allons préciser.

Contraintes de faisabilité. On se place dans le cas où le prix unitaire p de vente du bien équivaut au moins son coût c de fabrication c'est-à-dire $p \geq c$. Nous supposons que l'entreprise ne reçoit pas de subventions ou subsides si bien qu'elle doit nécessairement être rentable pour survivre.

Soit \underline{R} le revenu minimum de la société. Si pour $p \geq c$, $\underline{R} < R_0$, R_0 étant *le revenu minimal* R_0 donnant lieu à l'achat du bien X , alors il y aura nécessairement des personnes qui ne pourront pas avoir accès à ce bien. Dans le cas contraire, c'est-à-dire lorsque $R_0 \leq \underline{R}$, il est possible de donner accès à tous au bien X . Cependant, il n'est pas certain que l'entreprise décide d'une politique tarifaire telle que tous trouvent le bien accessible. On cherchera à identifier les conditions qui pousseraient l'entreprise à adopter une politique tarifaire inclusive.

3.4.4 Maximisation du profit et du bien-être

3.4.4.1 Maximisation du profit

On cherche à déterminer le prix optimal p qui répond à l'objectif de maximisation des profits du monopole. On considère l'équation (3.8).

Encadré 3.4.4.1 : Maximisation du Profit.

Le comportement maximisateur de profit de l'entreprise pousse cette dernière à l'efficacité économique contribuant ainsi à une meilleure allocation des ressources économiques rares. Dans un processus de production, la quantité d'inputs utilisés et l'output obtenu résultent de la maximisation du total des profits de l'entrepreneur conduisant ainsi à la maximisation du bien-être individuel et collectif, sous la contrainte que le prix du bien soit égal au coût marginal de production (Desai, 2008).

Pour déterminer le prix optimal p^* qui maximise le profit du monopole, on détermine et on annule l'équation obtenue issue des conditions de premier ordre (CPO) par rapport à p .

$$\begin{aligned}\frac{\partial \Pi(p)}{\partial p} &= \frac{\partial}{\partial p} \left(\int_{R_0}^{+\infty} (p - c) n(R) dR \right) \\ &= N - N(R_0) - (p - c) n(R_0) \frac{\partial R_0}{\partial p}\end{aligned}\quad (3.10)$$

En supposant l'existence d'une solution intérieure, le prix optimal p^* qui maximise le profit est tel que : (Annexe 3.A.1, équation 3.69)

$$\frac{p^* - c}{p^*} = \frac{N - N(R_0)}{n(R_0) R_0} \equiv \frac{1}{\varepsilon_{D/p}}. \quad (3.11)$$

L'expression de l'élasticité prix de la demande agrégée $\varepsilon_{D/p}$ est telle que ((Annexe 3.A.1, équation 3.70)

$$\varepsilon_{D/p} = \frac{n(R_0)}{N - N(R_0)} R_0. \quad (3.12)$$

Ce premier résultat n'est que la classique « règle de Lerner » $(p^* - c)/p^*$ (Lerner, 1934; Elzinga et Mills, 2011) selon laquelle, pour maximiser ses profits, une entreprise doit avoir une marge exactement égale à l'inverse de l'élasticité prix de la demande $(\varepsilon_{D/p})^{-1}$. Une conséquence directe de ce résultat est que, l'entreprise adoptera une marge d'autant plus élevée que la demande est inélastique. Dans notre cas de figure, ce résultat signifie tout naturellement que d'autant moins importante est la demande $(N - N(R_0))$ des ménages

ayant un minimum de revenu R_0 , d'autant plus important est le nombre de ménages (ceux ayant un revenu inférieur à R_0) qui n'achètent pas le bien. Autrement dit, l'augmentation du *revenu minimal* R_0 (*induite par celle de p_0*) *donnant accès au bien* fait diminuer la demande ($N - N(R_0)$) des ménages ayant accès au bien et fait augmenter du même coup le nombre de ménages qui n'achètent pas le bien. En conséquence, une condition suffisante pour avoir un niveau d'exclusion le plus faible que possible est que $N(R_0) \rightarrow N$. Ce qui renvoie à la problématique de la distribution des revenus dans cette société.

Prise en compte de la qualité

On se propose ici d'évaluer l'efficacité économique et la problématique de l'inclusion dans le cas où sont disponibles sur le marché deux qualités $q \in \{1, 2\}$ du bien X : une qualité supérieure (*qualité 2*) et une qualité inférieure (*qualité 1*). L'utilité associée à chaque qualité est telle que $u_2 > u_1 > 0$ avec des prix respectifs $p_2 > p_1 > 0$. On considère que les deux qualités du bien X sont produites ou bien par un monopole verticalement différencié ou bien par un duopole formé de deux firmes compétitives.

On reprend la fonction d'utilité telle que définie à l'équation (3.3)⁹ :

$$\mathcal{U}(1_x, R) = \ln(1 + u_q) + \ln(R - p_q)$$

où l'on note la qualité $q = 0$ quand aucune unité de X n'est achetée. On suppose aussi que $u_0 = 0$ et $p_0 = 0$.

Un ménage ayant un revenu $R > 0$ maximise son utilité en renonçant ou en acquérant une qualité unitaire du bien X . De manière formelle, le ménage préférera la qualité unitaire X_j à la qualité unitaire X_i ($i, j \neq i$) $\in \{0, 1, 2\}$ ² si et seulement si $\mathcal{U}(1_{X(j)}, R) > \mathcal{U}(1_{X(i)}, R)$ soit

$$\frac{R - p_j}{R - p_i} > \frac{1 + u_i}{1 + u_j}, \quad p_j > p_i.$$

9.

Le ménage préfère la qualité j à la qualité i si et seulement si

$$R \left(1 - \frac{1 + u_i}{1 + u_j} \right) > p_j - \left(\frac{1 + u_i}{1 + u_j} \right) p_i$$

c'est-à-dire quand

$$R \geq \frac{(1 + u_j)p_j - (1 + u_i)p_i}{u_j - u_i} \equiv R_{j,i}^c(p_j, p_i), \quad \forall (i, j \neq i) \in \{0, 1, 2\}^2. \quad (3.13)$$

où $R_{j,i}^c(p_j, p_i)$ désigne le revenu minimum du ménage préférant la qualité unitaire X_j à la qualité unitaire X_i . Comme démontré à l'annexe (...), pour les deux qualités disponibles, on doit avoir

$$R_{i,0}^c < R_{j,0}^c < R_{j,i}^c \quad (3.14)$$

et par conséquent il y aura à l'équilibre une demande non nulle si

$$p_j > \left(\frac{1 + 1/u_i}{1 + 1/u_j} \right) p_i. \quad (3.15)$$

Le ménage préfère renoncer au bien X quand $R < R_{1,0}^c$; il se procure la qualité q_1 tant que $R \in [R_{1,0}^c, R_{2,1}^c[$; il se procure la qualité q_2 tant que $R \geq R_{2,1}^c$. Formellement, on a :

$$D_1(p_2, p_1, R) = \mathbf{1}_{R \in [R_{1,0}^c, R_{2,1}^c[},$$

$$D_2(p_2, p_1, R) = \mathbf{1}_{R \in [R_{2,1}^c, +\infty[},$$

où $D_i(p_2, p_1, R)$, $i = 1, 2$ se réfère à la demande pour la qualité q_i , $i = 1, 2$. La demande unitaire pour les deux qualités est donc donnée par :

$$D(\mathbf{p}, R) = \sum_{i=1}^2 D_i(p_2, p_1, R) \in \{0, 1\}. \quad (3.16)$$

La demande agrégée pour chaque qualité (voir équation 3.10) s'écrit :

$$D_2(p_2, p_1) = \int_{R_{2,1}^c(p_2, p_1)}^{+\infty} D(p_2, p_1, R) g(R) dR = N - N(R_{2,1}^c(p_2, p_1)) \quad (3.17)$$

$$\begin{aligned} D_1(p_2, p_1) &= \int_{R_{1,0}^c(p_1)}^{R_{2,1}^c(p_2, p_1)} D(p_2, p_1, R) g(R) dR \\ &= N(R_{2,1}^c(p_2, p_1)) - N(R_{1,0}^c(p_1)). \end{aligned} \quad (3.18)$$

La demande agrégée pour les deux qualités est seulement une fonction du prix de la qualité q_1 et est telle que

$$D = D_2(p_2, p_1) + D_1(p_2, p_1) = N - N(R_{1,0}^c(p_1)) \equiv D(p_1).$$

Cas d'un monopole offrant deux produits verticalement différenciés. On suppose que la structure de marché n'a pas changé et que le monopole offre deux qualités du bien X . Les profits totaux du monopole s'écrivent :

$$\Pi(p_2, p_1) = (p_2 - c_2) D_2(p_2, p_1) + (p_1 - c_1) D_1(p_2, p_1). \quad (3.19)$$

Les conditions du premier ordre (*CPO*) associées sont données par :

$$0 = D_2(p_2, p_1) + (p_2 - c_2) \frac{\partial D_2(p_2, p_1)}{\partial p_2} + (p_1 - c_1) \frac{\partial D_1(p_2, p_1)}{\partial p_2}, \quad (3.20)$$

$$0 = D_1(p_2, p_1) + (p_2 - c_2) \frac{\partial D_2(p_2, p_1)}{\partial p_1} + (p_1 - c_1) \frac{\partial D_1(p_2, p_1)}{\partial p_1}. \quad (3.21)$$

Le monopole fixe le prix de chaque qualité en internalisant leur impact sur les deux qualités. En utilisant les équations (3.10), (3.17) et (3.18), on peut réécrire les équations (3.20) et (3.21) comme suit :

$$\begin{aligned}
0 &= N - N(R_{2,1}^c(p_2, p_1)) - (p_2 - c_2) n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_2} \\
&+ (p_1 - c_1) n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_2}, \tag{3.22}
\end{aligned}$$

$$\begin{aligned}
0 &= N(R_{2,1}^c(p_2, p_1)) - N(R_{1,0}^c(p_1)) - (p_2 - c_2) n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_1} \\
&+ (p_1 - c_1) \left[n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_1} - n(R_{1,0}^c(p_2, p_1)) \frac{\partial R_{1,0}^c(p_2, p_1)}{\partial p_1} \right]. \tag{3.23}
\end{aligned}$$

Pour déterminer une solution de ce système plus facile à interpréter, nous supposons que le revenu R du ménage suit une distribution de Pareto de type 1 de paramètre $\alpha > 0$. Le nombre de ménages $N(R)$ ayant au plus un revenu R (distribution cumulative) est définie par :

$$N(R) = N \left[1 - \left(\frac{R}{\underline{R}} \right)^\alpha \right] \quad \forall R \geq \underline{R}. \tag{3.24}$$

où \underline{R} désigne le revenu minimum de la société. La fonction de densité associée à $N(R)$ est donnée par :

$$n(R) = N \frac{\alpha}{R} \left(\frac{R}{\underline{R}} \right)^\alpha \quad \forall R \geq \underline{R}. \tag{3.25}$$

On réécrit alors les équations (3.22) et (3.23) de la manière suivante :

$$\begin{aligned}
0 &= N \left(\frac{R}{R_{2,1}^c} \right)^\alpha - (p_2 - c_2) \frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_2} + (p_1 - c_1) \frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_1} \\
0 &= N(\underline{R})^\alpha \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \right] - (p_2 - c_2) \frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_1} \\
&+ (p_1 - c_1) \left[\frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_1} - \frac{\alpha}{R_{1,0}^c} N \left(\frac{R}{R_{1,0}^c} \right)^\alpha \frac{\partial R_{1,0}^c}{\partial p_1} \right]
\end{aligned}$$

ou de manière équivalente

$$\begin{aligned}
0 &= N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \left\{ 1 - \frac{\alpha}{p_2} \frac{p_2}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_2} [(p_2 - c_2) - (p_1 - c_1)] \right\} \\
0 &= N (\underline{R})^\alpha \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \right] - \frac{\alpha}{p_1} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \left(\frac{p_1}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_1} \right) [(p_2 - c_2) - (p_1 - c_1)] \\
&\quad - \frac{\alpha}{p_1} N \left(\frac{R}{R_{1,0}^c} \right)^\alpha (p_1 - c_1) \left(\frac{p_1}{R_{1,0}^c} \frac{\partial R_{1,0}^c}{\partial p_1} \right)
\end{aligned}$$

Définissons les élasticités revenu suivantes :

$$\begin{aligned}
\eta_{R_{2,1}^c/p_2} &= \frac{p_2}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_2} = \frac{(1 + u_2) p_2}{(1 + u_2) p_2 - (1 + u_1) p_1} > 1 \\
\eta_{R_{2,1}^c/p_1} &= \frac{p_1}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_1} = \frac{-(1 + u_1) p_1}{(1 + u_2) p_2 - (1 + u_1) p_1} < 0, \\
\eta_{R_{1,0}^c/p_1} &= \frac{p_1}{R_{1,0}^c} \frac{\partial R_{1,0}^c}{\partial p_1} = 1.
\end{aligned} \tag{3.26}$$

On peut remarquer que $\eta_{R_{2,1}^c/p_2} + \eta_{R_{2,1}^c/p_1} = 1$. Ainsi, le système se réécrit comme suit :

$$0 = N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \left\{ 1 - \alpha \eta_{R_{2,1}^c/p_2} \left[\frac{(p_2 - c_2) - (p_1 - c_1)}{p_2} \right] \right\} \tag{3.27}$$

$$\begin{aligned}
0 &= N (\underline{R})^\alpha \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \right] - \alpha N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \eta_{R_{2,1}^c/p_1} \left[\frac{(p_2 - c_2) - (p_1 - c_1)}{p_1} \right] \\
&\quad - \alpha N \left(\frac{R}{R_{1,0}^c} \right)^\alpha \left(\frac{p_1 - c_1}{p_1} \right) \eta_{R_{1,0}^c/p_1}.
\end{aligned} \tag{3.28}$$

Remarquons que l'élasticité-prix de la demande agrégée ϵ_{D_2/p_2} pour la qualité supérieure est telle que

$$\begin{aligned}
\epsilon_{D_2/p_2} &= \frac{p_2}{D_2} \frac{\partial D_2}{\partial p_2} = \frac{p_2}{N - N(R_{2,1}^c(p_2, p_1))} n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_2} \\
&= p_2 N^{-1} \left(\frac{R}{R_{2,1}^c} \right)^{-\alpha} \frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_2} \\
&= \alpha \frac{p_2}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_2} \\
&= \alpha \eta_{R_{2,1}^c/p_2} > \alpha, \quad \forall \alpha > 1.
\end{aligned}$$

Ce qui signifie que la demande pour la qualité supérieure est plus élastique ($\epsilon_{D_2/p_2} > \alpha > 1$)¹⁰ du fait de la disponibilité de la qualité inférieure. De l'équation (3.27), on en déduit :

$$\begin{aligned} \frac{(p_2 - c_2) - (p_1 - c_1)}{p_2} &= \frac{1}{\alpha \eta_{R_{2,1}^c/p_2}}. \\ &= \frac{1}{\epsilon_{D_2/p_2}} < \frac{1}{\alpha} \end{aligned} \quad (3.29)$$

et donc, en présence d'une qualité inférieure, le gain net en proportion du prix chargé pour la qualité supérieure est inférieur à celui qu'on aurait observé dans le cas de la disponibilité d'une seule qualité.

On peut alors récrire l'équation (3.28) comme suit :

$$\begin{aligned} 0 &= N(\underline{R})^\alpha \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \right] - N\left(\frac{\underline{R}}{R_{2,1}^c}\right)^\alpha \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} \\ &\quad - \alpha N\left(\frac{\underline{R}}{R_{1,0}^c}\right)^\alpha \left(\frac{p_1 - c_1}{p_1}\right), \quad \eta_{R_{1,0}^c/p_1} = 1 \\ &= N(\underline{R})^\alpha \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} - \alpha (R_{1,0}^c)^{-\alpha} \left(\frac{p_1 - c_1}{p_1}\right) \right] \\ &= (R_{1,0}^c)^{-\alpha} \left[1 - \alpha \left(\frac{p_1 - c_1}{p_1}\right) \right] - (R_{2,1}^c)^{-\alpha} \left(1 + \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} \right) \\ &= 1 - \alpha \left(\frac{p_1 - c_1}{p_1}\right) - \left(1 + \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} \right) \left(\frac{R_{1,0}^c}{R_{2,1}^c}\right)^\alpha. \end{aligned}$$

10. On considère ici les valeurs de $\alpha > 1$ en raison de la loi de *Van Der Wijk* qui stipule que « si l'on considère un niveau de revenu de base R , le revenu moyen $z(R)$ de tout sous-groupe d'individus ayant un revenu au moins aussi élevé que le revenu de base est égal à kR où $k = \alpha/(\alpha - 1)$ ». Cette loi se vérifie pour la distribution de Pareto de type 1. En effet,

$$\begin{aligned} z(R) &= E(R_p/R_p \geq R), \quad R_p \sim \mathcal{L}_P(\alpha, R) \\ &\stackrel{\text{déf}}{=} \frac{\int_R^{+\infty} y n(y) dy}{\int_R^{+\infty} n(y) dy} = \frac{\int_R^{+\infty} \frac{dy}{y^\alpha}}{\int_R^{+\infty} \frac{dy}{y^{\alpha+1}}}, \quad Y \sim \mathcal{L}_P(\alpha, y) \\ &= \left(\frac{\alpha}{\alpha - 1}\right) R, \quad \text{pour tout } \alpha > 1 \\ &= kR, \quad k = \frac{\alpha}{\alpha - 1}. \end{aligned}$$

Par suite, la marge pour la qualité inférieure est donnée par :

$$\begin{aligned} \frac{p_1 - c_1}{p_1} &= \left[1 - \left(1 + \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} \right) \left(\frac{R_{1,0}^c}{R_{2,1}^c} \right)^\alpha \right] \frac{1}{\alpha} \\ &= \left[1 - \left(\frac{u_2 - u_1}{1 + u_2} \right) \left(\frac{R_{1,0}^c}{R_{2,1}^c} \right)^\alpha \right] \frac{1}{\alpha} < \frac{1}{\alpha}. \end{aligned} \quad (3.30)$$

car

$$\begin{aligned} 1 + \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} &= 1 - \frac{1 + u_1}{1 + u_2} \\ &= \frac{u_2 - u_1}{1 + u_2} < 1 \end{aligned}$$

Ce qui signifie que la marge obtenue pour la qualité inférieure du bien X est plus petite que la marge qu'on aurait obtenue en absence de la qualité supérieure.

De l'équation (3.29), la marge pour la qualité supérieure est telle que

$$\begin{aligned} \frac{p_2 - c_2}{p_2} &= \frac{1}{\alpha \eta_{R_{2,1}^c/p_2}} + \frac{p_1}{p_2} \frac{p_1 - c_1}{p_1} \\ &= \left\{ \frac{1}{\eta_{R_{2,1}^c/p_2}} + \frac{p_1}{p_2} \left[1 - \left(1 + \frac{\eta_{R_{2,1}^c/p_1} p_2}{\eta_{R_{2,1}^c/p_2} p_1} \right) \left(\frac{R_{1,0}^c}{R_{2,1}^c} \right)^\alpha \right] \right\} \frac{1}{\alpha} \\ &= \left\{ \frac{1}{\eta_{R_{2,1}^c/p_2}} + \frac{p_1}{p_2} \left[1 - \left(\frac{u_2 - u_1}{1 + u_2} \right) \left(\frac{R_{1,0}^c}{R_{2,1}^c} \right)^\alpha \right] \right\} \frac{1}{\alpha} < \frac{1}{\alpha}. \end{aligned} \quad (3.31)$$

Les marges obtenues aux équations (3.30) et (3.31) permettent d'observer que la disponibilité de plusieurs qualités augmente l'accès des ménages au bien. De plus, d'autant plus égalitaire est la distribution du revenu, d'autant plus que l'entreprise tend à adopter une tarification à la marge. Formellement, on a :

$$\lim_{\alpha \rightarrow +\infty} p_1 = c_1 \quad \text{et} \quad \lim_{\alpha \rightarrow +\infty} p_2 = c_2.$$

Cas d'un duopole. On considère un duopole formé de deux firmes compétitives produisant l'une la qualité supérieure du bien X et l'autre la qualité inférieure. Les profits

totaux du duopole s'écrivent comme ceux du monopole définis à l'équation (3.19). Par contre, les *CPOs* dans le cas du duopole diffèrent de celles dans le cas du monopole du fait que chaque firme ne tient pas compte de la politique tarifaire menée par l'autre de sorte que :

$$0 = D_2(p_2, p_1) + (p_2 - c_2) \frac{\partial D_2(p_2, p_1)}{\partial p_2}, \quad (3.32)$$

$$0 = D_1(p_2, p_1) + (p_1 - c_1) \frac{\partial D_1(p_2, p_1)}{\partial p_1}. \quad (3.33)$$

En utilisant les résultats obtenus aux équations (3.10), (3.17) et (3.18), on peut récrire les équations (3.32) et (3.33) comme suit :

$$0 = N - N(R_{2,1}^c(p_2, p_1)) - (p_2 - c_2) n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_2} \quad (3.34)$$

$$0 = N(R_{2,1}^c(p_2, p_1)) - N(R_{1,0}^c(p_1)) + (p_1 - c_1) \left[n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_1} - n(R_{1,0}^c(p_2, p_1)) \frac{\partial R_{1,0}^c(p_2, p_1)}{\partial p_1} \right]. \quad (3.35)$$

En utilisant les fonctions $N(R)$ et $n(R)$ définies aux équations (3.24) et (3.25), l'équation (3.34) peut se récrire de la façon suivante :

$$\begin{aligned} 0 &= N - N \left[1 - \left(\frac{R}{R_{2,1}^c} \right)^\alpha \right] - (p_2 - c_2) \frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_2} \\ &= N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \left[1 - \alpha \left(\frac{p_2 - c_2}{p_2} \right) \frac{p_2}{R_{2,1}^c} \frac{\partial R_{2,1}^c}{\partial p_2} \right] \\ &= N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \left[1 - \alpha \left(\frac{p_2 - c_2}{p_2} \right) \eta_{R_{2,1}^c/p_2} \right]. \end{aligned}$$

Ainsi,

$$\frac{p_2 - c_2}{p_2} = \frac{1}{\alpha \eta_{R_{2,1}^c/p_2}}. \quad (3.36)$$

En suivant la même démarche pour l'équation (3.35), on obtient :

$$\begin{aligned}
0 &= N(R_{2,1}^c(p_2, p_1)) - N(R_{1,0}^c(p_1)) \\
&+ (p_1 - c_1) \left[n(R_{2,1}^c(p_2, p_1)) \frac{\partial R_{2,1}^c(p_2, p_1)}{\partial p_1} - n(R_{1,0}^c(p_2, p_1)) \frac{\partial R_{1,0}^c(p_2, p_1)}{\partial p_1} \right] \\
&= N \left[1 - \left(\frac{R}{R_{2,1}^c} \right)^\alpha \right] - N \left[1 - \left(\frac{R}{R_{1,0}^c} \right)^\alpha \right] \\
&+ (p_1 - c_1) \left[\frac{\alpha}{R_{2,1}^c} N \left(\frac{R}{R_{2,1}^c} \right)^\alpha \frac{\partial R_{2,1}^c}{\partial p_1} - \frac{\alpha}{R_{1,0}^c} N \left(\frac{R}{R_{1,0}^c} \right)^\alpha \frac{\partial R_{1,0}^c}{\partial p_1} \right] \\
&= N(\underline{R})^\alpha \left\{ \left[(R_{1,0}^c)^{-\alpha} - (R_{2,1}^c)^{-\alpha} \right] + \right. \\
&\left. + \left(\frac{p_1 - c_1}{p_1} \right) \alpha \left[\eta_{R_{2,1}^c/p_1} (R_{2,1}^c)^{-\alpha} - \eta_{R_{1,0}^c/p_1} (R_{1,0}^c)^{-\alpha} \right] \right\} \\
&= N(\underline{R})^\alpha \left\{ (R_{1,0}^c)^{-\alpha} \left[1 - \left(\frac{p_1 - c_1}{p_1} \right) \alpha \right] - (R_{2,1}^c)^{-\alpha} \left[1 - \left(\frac{p_1 - c_1}{p_1} \right) \alpha \eta_{R_{2,1}^c/p_1} \right] \right\}.
\end{aligned}$$

Par suite,

$$\begin{aligned}
\frac{p_1 - c_1}{p_1} &= \frac{1}{\alpha} \frac{(R_{2,1}^c)^{-\alpha} - (R_{1,0}^c)^{-\alpha}}{\left[\eta_{R_{2,1}^c/p_1} (R_{2,1}^c)^{-\alpha} - (R_{1,0}^c)^{-\alpha} \right]}, \quad \eta_{R_{1,0}^c/p_1} = 1 \\
&= \frac{1}{\alpha} \left[\frac{(R_{1,0}^c)^\alpha - (R_{2,1}^c)^\alpha}{\eta_{R_{2,1}^c/p_1} (R_{1,0}^c)^\alpha - (R_{2,1}^c)^\alpha} \right] \\
&= \frac{1}{\alpha} \left[\frac{(R_{2,1}^c)^\alpha - (R_{1,0}^c)^\alpha}{(R_{2,1}^c)^\alpha - \eta_{R_{2,1}^c/p_1} (R_{1,0}^c)^\alpha} \right] \tag{3.37}
\end{aligned}$$

$$= \frac{1}{\alpha} \left[\frac{1 - (R_{1,0}^c/R_{2,1}^c)^\alpha}{1 - \eta_{R_{2,1}^c/p_1} (R_{1,0}^c/R_{2,1}^c)^\alpha} \right], \quad \eta_{R_{2,1}^c/p_1} < 0. \tag{3.38}$$

Il apparait clairement que les marges obtenues aux équations (3.36) et (3.38) sont inférieures à la marge $1/\alpha$ obtenue si une seule qualité était disponible. On peut aussi faire remarquer que

$$\lim_{\alpha \rightarrow +\infty} p_1 = c_1 \quad \text{et} \quad \lim_{\alpha \rightarrow +\infty} p_2 = c_2.$$

Autrement dit, d'autant plus égalitaire est la distribution de revenu, d'autant plus que le prix chargé par l'entreprise se rapproche du coût marginal.

3.4.4.2 Maximisation du bien-être

La maximisation des profits (surplus) des producteurs résulte de l'efficacité économique de l'entreprise. Dans un système de marché, cette efficacité économique doit correspondre à la maximisation du surplus (bien-être) social (collectif) c'est-à-dire non seulement à la maximisation du surplus des producteurs mais également à celle des consommateurs. L'idée est déterminer un niveau de prix optimal du point de la société.

Pour ce faire, on considère l'équation (3.9) définissant le bien-être social. La *CPO* en p est telle que

$$\begin{aligned} \frac{\partial W(p)}{\partial p} &= \frac{\partial}{\partial p} \left(\left\{ \int_0^{R_0} \ln(R) n(R) dR + \int_{R_0}^{+\infty} [\ln(1+u) + \ln(R-p)] n(R) dR \right\} \right. \\ &\quad \left. + \left\{ \int_{R_0}^{+\infty} (p-c) n(R) dR \right\} \right) \\ &= -(p-c) n(R_0) \frac{\partial R_0}{\partial p} + \int_{R_0}^{+\infty} \left[1 - \frac{1}{R-p} \right] n(R) dR. \end{aligned} \quad (3.39)$$

Encadré 3.4.4.2 : Maximisation du bien-être social.

L'optimum social, résultat de la maximisation du bien-être social, est décrit et caractérisé à travers les deux théorèmes du bien-être (Baujard, 2013b). Le premier théorème stipule que « les équilibres concurrentiels sont optimaux au sens de Pareto si les préférences individuelles sont monotones et si les marchés sont complets ^a ». Le deuxième théorème établit qu'« une allocation optimale au sens de Pareto peut être atteinte dans le cadre d'un équilibre concurrentiel si les dotations initiales sont redistribuées de manière appropriée ». Nous nous plaçons dans ce cadre de maximisation du bien-être social pour déterminer le prix socialement optimal.

^a. On dit qu'un marché est complet si et seulement si on peut générer par au moins un portefeuille n'importe quel profit de revenu.

Vérifions que le prix optimal p qui maximise le profit ne maximise pas le bien-être social. En effet, à partir de la *CPO* de l'équation (3.10) de profit, quand le prix choisi pour maximiser le profit, on déduit : $N - N(R_0) = (p-c) n(R_0) \frac{\partial R_0}{\partial p}$ et donc (Annexe

3.A.3, équation 3.71)

$$\begin{aligned} \frac{\partial W(p)}{\partial p} &= -[N - N(R_0)] + \int_{R_0}^{+\infty} \left[1 - \frac{1}{R-p}\right] n(R) dR \\ &= - \int_{R_0}^{+\infty} \left[\frac{1}{R-p}\right] n(R) dR < 0, \quad \int_{R_0}^{+\infty} n(R) dR = N - N(R_0). \end{aligned} \quad (3.40)$$

Ce résultat signifie tout simplement que le prix qui maximise le profit est trop élevé et donc la marge qui maximise le bien-être est inférieure à celle qui maximise le profit.

Quelques manipulations (Annexe 3.A.3, équation 3.74) permettent de déterminer le prix socialement optimal p^* définissant la marge qui maximise le bien-être social. L'expression de $\varepsilon_{D/p}$ obtenue à l'équation 3.12 permet de simplifier l'équation en annulant la CPO déterminée en (3.39) pour obtenir :

$$\frac{p^* - c}{p^*} = \frac{\int_{R_0^*}^{+\infty} \left[1 - \frac{1}{R-p^*}\right] n(R) dR}{p^* n(R_0^*) \frac{\partial R_0^*}{\partial p}} \quad (3.41)$$

$$= \frac{1 - \int_{R_0^*}^{+\infty} \left(\frac{1}{R-p^*}\right) \frac{n(R)}{N-N(R_0^*)} dR}{\varepsilon_{D/p}}. \quad (3.42)$$

L'expression $(R - p^*)^{-1}$ ¹¹ correspond à l'utilité marginale du revenu. est donc une fonction de la moyenne des utilités marginales du revenu *des personnes qui achètent le bien* (et pas

11. On peut bien remarquer que la fonction d'utilité indirecte, formellement définie par :

$$\mathcal{V}(p, R) = \max_{x, R} \{ \mathcal{U}(x, R) / px \geq R, x \geq 0 \}, \quad p \geq 0, R \geq 0$$

est, dans ce cas, tout simplement donnée par :

$$\mathcal{V}(p, R) = \ln(1 + u \cdot I_{x \geq 1}) + \ln(R - p).$$

Par suite, l'utilité marginale du revenu correspond à

$$\frac{d\mathcal{V}}{dR} = \frac{1}{R-p}.$$

De plus, comme $R > R_0$ et donc

$$\frac{1}{R-p} < \frac{1}{R_0-p} = \frac{u}{p}.$$

Par suite

$$\frac{d\mathcal{V}}{dR} < \frac{1}{R_0-p} = \frac{u}{p}$$

de toute la population) $\int_{R_0}^{+\infty} \left[\frac{1}{R-p} \right] \frac{n(R)}{N-N(R_0)} dR$. Il en résulte que, si cette utilité marginale moyenne est supérieure à 1¹², alors la maximisation du bien-être social requiert *a priori* de 'subventionner' les consommateurs c'est-à-dire d'adopter un prix strictement inférieur au coût marginal c . Si l'on opère sous contrainte de rentabilité, on fixera alors le prix au niveau le plus bas compatible avec une telle contrainte, c'est-à-dire au coût marginal : $p^* = c$.

Il est à noter que, contrairement au résultat classique obtenu dans le cadre des utilités quasi-linéaires, si l'utilité marginale moyenne du revenu des ménages qui consomment le bien est inférieure à 1, alors le prix socialement optimal du bien est strictement supérieur au coût marginal. Autrement dit, la tarification au coût marginal n'est socialement efficace que si l'on suppose que l'utilité marginale du revenu est constante.

12. En référence au type de préférences adoptées (voir 3.4.1) pour développer ce modèle, une condition suffisante pour avoir une marge positive est que l'utilité retirée de la consommation d'une unité du bien soit au plus égale à son prix p . La marge est nulle (efficacité économique) pour les $n(R_0)$ ménages indifférents entre acheter ou ne pas acheter le bien c'est-à-dire les ménages ayant exactement un niveau de revenu R_0 . Ce qui peut se vérifier très facilement. En effet, comme $R > R_0$ alors

$$\frac{1}{R-p} < \frac{1}{R_0-p} = \frac{u}{p}.$$

Par suite,

$$\int_{R_0}^{+\infty} \left(\frac{1}{R-p} \right) \frac{n(R)}{N-N(R_0)} dR < \frac{u}{p} \int_{R_0}^{+\infty} \frac{n(R)}{N-N(R_0)} = \frac{u}{p}$$

ou de manière équivalente

$$1 - \int_{R_0}^{+\infty} \left(\frac{1}{R-p} \right) \frac{n(R)}{N-N(R_0)} dR > 1 - \frac{u}{p}.$$

Par conséquent

$$\frac{1 - \int_{R_0}^{+\infty} \left(\frac{1}{R-p} \right) \frac{n(R)}{N-N(R_0)} dR}{\varepsilon_{D/p}} > \frac{1 - \frac{u}{p}}{\varepsilon_{D/p}}.$$

Ainsi, une condition suffisante pour que la marge soit positive ou nulle est que $\left(1 - \frac{u}{p}\right) \geq 0$ c'est-à-dire $p \leq u$. Il en résulte que $\int \left(\frac{1}{R-p} \right) \frac{n(R)}{N-N(R_0)} dR \leq 1$.

3.5 Exclusion

Comme il a été précisé au niveau de l'introduction, nous sommes intéressés à l'analyse de l'importance du nombre de ménages exclus de la consommation du bien en fonction des caractéristiques de la distribution des revenus au niveau de la société. Cette analyse se porte et sur le comportement de l'entreprise qui maximise son profit et du point de vue de la maximisation du bien-être social.

Tout au long de cet essai, les concepts « accès » et « exclusion » constituent deux des principaux concepts fréquemment utilisés.

Le concept « exclusion » par opposition à celui d'« inclusion », pris en général dans le sens d'exclusion sociale, est caractérisé de manière générale, par un ensemble « de processus et de situations d'exclusion des individus et des groupes, des opportunités, des bénéfices et des droits qui sont communément disponibles dans les sociétés contemporaines ».

Encadré 3.5 : Le concept "exclusion".

Selon Silver (1994), l'exclusion se définit par opposition à différentes notions contextuelles d'intégration, de solidarité et de citoyenneté. De manière générale, le terme « exclusion » par opposition à celui d'« inclusion », pris en général dans le sens d'exclusion sociale, est caractérisé par un ensemble « de processus et de situations d'exclusion des individus et des groupes, des opportunités, des bénéfices et des droits qui sont communément disponibles dans les sociétés contemporaines » (Trbanc *et al.*, 1996). Selon Hamilton (2009), le renforcement des inégalités induit un accroissement de l'écart entre les hauts et les bas revenus contribuant du même coup à *exclure* les individus à faibles revenus - ce que nous nous proposons de modéliser - de la consommation d'un certain nombre de biens et services. Les gens à bas revenus sont alors qualifiés « de matériellement déshérités, - c'est-à-dire incapables de se procurer les biens et services répondant à un standard social de vie acceptable - ou socialement déshérités - ils n'arrivent pas à vivre ce que la vie offre et sont du même coup privés de la joie d'une vie heureuse. En raison de leur incapacité à correspondre aux normes de consommation en hausse qui sont nécessaires à un niveau de vie socialement acceptable, les consommateurs pauvres sont donc stigmatisés».

L'accès¹³ est défini ici comme la « capacité¹⁴ » à tirer des bénéfices d'un ensemble de « choses » (ressources) ou comme un ensemble de moyens dont dispose une personne pour tirer des bénéfices d'un ensemble de choses.

13. Du point de vue de l'économie du bien-être, la problématique de l'accès n'est pas tellement traitée dans la littérature en dépit de la justification exprimée à travers des programmes de gouvernement pour un accès égal à certains biens et services. L'une des rares recherches (Olsen et Rogers, 1991) à se pencher sur cette question, l'a abordée en termes *d'accès égal* à un bien dans une situation d'intervention ou de non intervention de l'état et où les agents valorisent la notion d'accès égal. Ces chercheurs argüent par définition que les gens ont un accès égal à un bien si et seulement si ils ont la capacité de consommer la même quantité du bien. Ils ont caractérisé cette définition à travers cinq propositions dans un modèle qu'ils ont développé pour deux agents en situation d'intervention ou de non intervention de l'état. Les deux premières propositions ont mis en lumière l'efficacité d'une intervention de l'état dans le cas où les agents valorisent la notion d'accès égal. La première proposition, en supposant qu'il existe une politique de l'état capable d'améliorer le sort des deux, stipule que : « si les deux individus valorisent la notion d'égal accès et si la situation sans intervention de l'état implique un accès inégal, alors cette situation est inefficace ». Ensuite, en supposant que quelqu'un sera pire si le gouvernement poursuit une politique qui augmente l'inégalité de l'accès au premier ou le laisse inchangé, la deuxième proposition pointe donc l'importance et l'efficacité de l'intervention de l'état. Cette proposition postule que « si les deux individus valorisent la notion d'égal accès et si la situation sans intervention de l'état implique un accès inégal, tous les états préférés à la situation initiale impliquent moins d'inégalité d'accès au bien ».

Une analyse du contexte de l'emploi du concept « accès », de sa définition, de ses différents usages, de son contenu réel et les modalités de son opérationnalisation peut être trouvée dans l'article publié par Ndour et Boidin (2012).

14. C'est-à-dire l'habileté de certains acteurs à influencer sur les pratiques et les idées des autres

Encadré 3.5 : Le concept "accès".

Dans un article publié par Ribot et Peluso (2003), le concept « accès » est défini comme la « capacité ^a » à tirer des bénéfices d'un ensemble de « choses » (ressources) ou comme un ensemble de moyens dont dispose une personne pour tirer des bénéfices d'un ensemble de choses. L'accès est influençable à travers un ensemble de mécanismes structurel et relationnel qui façonnent comment les bénéfices sont gagnés, contrôlés et maintenus. Ces mécanismes agissent principalement par le biais de l'autorité, des identités, de la connaissance, de la technologie, du capital, du marché, du travail et des relations sociales. L'accès tel que nous le concevons dans le cadre de cet essai, c'est donc un *processus* qui amène les membres d'une communauté à passer d'une dotation initiale à une acquisition de ressources (actifs) reconnue et de capacités (Ribot et Peluso, 2003; Geiser *et al.*, 2011). Penchansky et Thomas (1981) ont décliné le concept d'accès en cinq dimensions objectivement opérationnalisables. Selon ces chercheurs, ces cinq dimensions regroupent les aspects tels : *i) la disponibilité* tenant compte des besoins du consommateur ; *ii) l'accessibilité* faisant référence à la localisation du bien par rapport à la localisation du consommateur ; *iii) l'accommodation* ayant trait à la façon dont les ressources mises à la disposition des consommateurs sont organisées et comment ces derniers perçoivent ces ressources et leur organisation, les approprient et les utilisent ; *iv) l'abordabilité* caractérisée par le rapport qualité/prix eu égard aux ressources (revenu) du consommateur client et enfin *v) l'acceptabilité* jaugée à travers la relation entre les attitudes des consommateurs en rapport avec les caractéristiques individuelles et pratiques des fournisseurs et les attitudes des fournisseurs en rapport avec les caractéristiques individuelles des consommateurs.

a. C'est-à-dire l'habilité de certains acteurs à influencer sur les pratiques et les idées des autres

3.5.1 Maximisation du profit et exclusion

Étant donné le prix maximisant les profits du monopole, les ménages qui ont un revenu inférieur à R_0 n'ont pas un revenu suffisant pour accéder au bien X . Soit \underline{R} le revenu minimum de la société. La valeur seuil en \underline{R} du prix p définie à partir de l'inégalité (3.4) devient $\underline{p} = [u/(1+u)] \underline{R}$. Si le prix du bien dépasse cette valeur seuil alors $R_0 > \underline{R}$ et la population exclue s'élève à :

$$N_0 \equiv N(R_0) = \int_0^{R_0} n(R) dR.$$

Si le coût c est trop élevé, il n'est pas possible que le bien X soit produit et vendu de manière rentable sans qu'il n'y ait exclusion. Par contre, si

$$R_0(c) = \frac{1+u}{u}c \leq \underline{R}, \quad (3.43)$$

il est parfaitement possible de produire le bien X sans qu'il n'y ait exclusion.

Supposons qu'il soit possible de produire le bien X de manière rentable sans pour autant exclure personne. Autrement dit, supposons que la condition (3.43) soit vérifiée. On peut alors définir le prix

$$\underline{p} = \frac{u}{1+u}R \quad (3.44)$$

qui est la disposition à payer pour le bien X des ménages aux plus bas revenus. C'est aussi bien-sûr le prix en deçà duquel il n'y a pas d'exclusion.

Par hypothèse $\underline{p} \geq c$. La question que nous posons ici est celle de savoir en quelles circonstances le prix p^* qui maximise le profit vérifie $p^* \leq \underline{p}$.

Plus précisément, on veut identifier les conditions sur la distributions des revenus $n(R)$ qui sont telles que le prix défini en (3.11) vérifie $p^* < \underline{p}$ – si la solution du programme de maximisation des profits est « intérieure » – et de manière plus générale – lorsque techniquement parlant, la maximisation du profit donne lieu à une solution en coin – que l'entreprise n'a jamais intérêt à adopter un prix $p^* > \underline{p}$.

3.5.2 Maximisation du bien-être social et exclusion

Nous avons pu établir en (3.40) que le prix optimal maximisant le profit est trop élevé comparativement au prix maximisant le bien-être social. Néanmoins, le prix maximisant le bien-être social garantit-il forcément un accès au bien à tous les ménages ayant un revenu $R < R_0$ qui le souhaitent? Autrement dit, la maximisation du bien-être social est-elle équivalente à la minimisation de l'exclusion? De ce fait, on se ramène à la même problématique élaborée dans le cadre de la maximisation du profit et exclusion.

3.5.3 Illustrations des résultats

Dans cette sous-section, sont adoptées deux des distributions de probabilités bien connues pour leur utilisation dans la modélisation des revenus pour illustrer les résultats théoriques. Il s'agit de la distribution de Pareto et de la distribution Log-normale.

3.5.3.1 Le cas de la distribution de Pareto

La distribution de Pareto qui est utilisée est celle de type 1¹⁵. C'est une des distributions de probabilité abondamment utilisée dans la littérature pour modéliser la distribution des revenus.

Distribution de Pareto de type 1

1. Définition. Soit R_p le montant de revenus détenu par un individu quelconque d'une population donnée. La variable aléatoire R_p suit une loi de Pareto de type 1 de paramètres (α, \underline{R}) si ses fonctions de distribution (fonction de répartition) et de densité sont respectivement définies par :

$$F_{R_p}(R) = 1 - \left(\frac{R}{\underline{R}}\right)^\alpha, \quad \forall R \geq \underline{R}, \text{ et } \alpha > 0 \quad (3.45)$$

$$f_{R_p}(R) = \alpha \left(\frac{\underline{R}^\alpha}{R^{\alpha+1}}\right), \quad \forall R \geq \underline{R} \text{ et } \alpha > 0.$$

où \underline{R} est le revenu minimum de la population. On note : $R_p \sim \mathcal{L}_P(\alpha, \underline{R})$. La fonction $F_{R_p}(R)$ ¹⁶ détermine « la proportion d'individus ayant un revenu inférieur à R ».

15. Dans la littérature, on trouve jusqu'à quatre types de distributions de Pareto. La distribution que nous utilisons est type 1. Sa fonction de répartition $F_X(x)$ est définie par : $F_X(x) = 1 - \left(\frac{x_m}{x}\right)^\alpha$, $\forall x \geq x_m$. Par contre, si on remplace $\left(\frac{x_m}{x}\right)$ par $\left(1 + \frac{x_m}{x-\mu}\right)$, $\forall x > \mu$ pour $\alpha, x_m > 0$ et $\mu \in \mathbb{R}$, la distribution de Pareto est de type 2. La distribution de Pareto est dite de type 3 si $F_X(x) = 1 - \left[1 + \left(\frac{x_m}{x-\mu}\right)^{-1/\gamma}\right]^{-1}$, $\forall x > \mu$ pour $\gamma, x_m > 0$ et $\mu \in \mathbb{R}$. Enfin, La distribution de Pareto est de type 4 si $F_X(x) = 1 - \left[1 + \left(\frac{x_m}{x-\mu}\right)^{-1/\gamma}\right]^{-\alpha}$, $\forall x > \mu$ pour $\alpha, \gamma, x_m > 0$ et $\mu \in \mathbb{R}$ (Arnold, 2014; Fournier et Piketty, 2015a).

16. L'indice G d'inégalité de Gini associé à cette distribution a pour expression :

$$G = \frac{1}{2\alpha - 1}, \quad \alpha > \frac{1}{2}.$$

Un autre indice d'inégalité intéressant bien connu est l'indice P_X de Pietra ou de Schutz. Il permet de déterminer la proportion du total de revenus nécessaires à ré-allouer pour atteindre une répartition de revenus égalitaire. Cet indice a pour expression :

$$P_X = \frac{(\alpha - 1)^{\alpha-1}}{\alpha^\alpha}, \quad \alpha > 1.$$

2. Les fonctions de survie $S(R)$ et fonction de hasard $h(R)$ ¹⁷ sont données par :

$$\begin{aligned} S(R) &\equiv P(R_p > R) & (3.46) \\ &= 1 - F_{R_p}(R) \\ &= \left(\frac{R}{R_0}\right)^\alpha \end{aligned}$$

$$\begin{aligned} h(R) &\equiv \frac{f_{R_p}(R)}{S(R)} & (3.47) \\ &= \frac{\alpha}{R} \end{aligned}$$

et donc $Rh(R) = \alpha$.

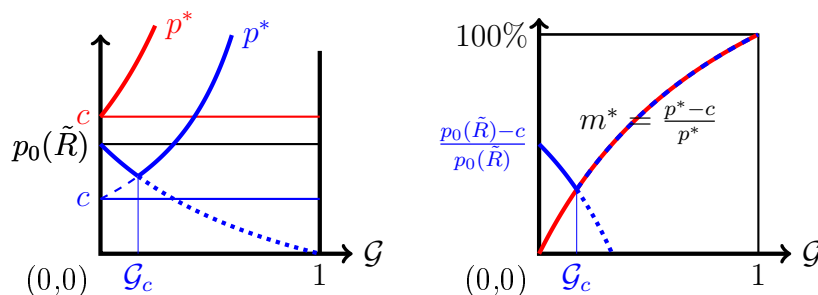
De même, de l'équation (3.12), l'expression de l'élasticité-prix de la demande agrégée est donnée par :

$$\begin{aligned} \varepsilon_{D/p} &= \frac{n(R_0)}{N - N(R_0)} R_0 \implies \\ \varepsilon_{D/p} &= \alpha \left(\frac{R}{R_0^{\alpha+1}} \right) \frac{1}{\left(\frac{R}{R_0}\right)^\alpha} R_0 \\ &= \alpha. \end{aligned}$$

et donc la règle de Lerner devient :

$$\frac{p^* - c}{p^*} = \frac{1}{\alpha}.$$

17. Les fonctions de survie et de hasard sont empruntées à la démographie. La fonction de survie définit ici la proportion de ménages ayant au moins un niveau de revenu R . La fonction du hasard définit alors la proportion de ménages ayant un revenu R parmi ceux ayant au moins ce revenu.

FIGURE 3.5.1 – Prix p^* et marge m^* du monopole en fonction de l'indice de Gini \mathcal{G} .**Encadré 3.5.3.1 : Distribution de Pareto.**

Les raisons principales évoquées pour l'utilisation de la distribution de Pareto sont les suivantes (Cowell, 2011; Lubrano, 2013) :

- La distribution de Pareto de type 1 est définie en fonction d'un seul paramètre facile à interpréter, facilitant ainsi le calcul et l'interprétation de différentes mesures d'inégalité.
- Cette distribution de Pareto vérifie la loi de Van Der Wijk ^a
- Les courbes de Lorenz dérivées de cette distribution ne se coupent pas. Ce qui facilite la comparaison du niveau d'inégalité pour deux distributions de revenus à moyenne égale à partir de deux courbes de Lorenz pour des valeurs différentes de α .
- Une propriété des queues ^b de cette distribution qui la rend intéressante pour un test empirique.

a. En résumé, la loi de Van der Wijk stipule que « si l'on considère un niveau de revenu de base y , le revenu moyen de tout sous-groupe d'individus ayant un revenu au moins aussi élevé que le revenu de base est égal à ky où $k = \alpha / (\alpha - 1)$ ». En effet, pour tout niveau de revenu R , le revenu moyen $z(R)$ du sous-groupe qui gagne au moins un revenu R est égal à :

$$\begin{aligned}
 z(R) &= E(R_p / R_p > R), \quad R_p \sim \mathcal{L}_P(\alpha, \underline{R}) \\
 &\stackrel{\text{déf}}{=} \frac{\int_R^{+\infty} yn(y) dy}{\int_R^{+\infty} n(y) dy} = \frac{\int_R^{+\infty} \frac{dy}{y^\alpha}}{\int_R^{+\infty} \frac{dy}{y^{\alpha+1}}}, \quad Y \sim \mathcal{L}_P(\alpha, y) \\
 &= \left(\frac{\alpha}{\alpha - 1} \right) R \\
 &= kR, \quad k = \frac{\alpha}{\alpha - 1}
 \end{aligned}$$

b. Dans un graphique *log-log*, la queue de la distribution de Pareto de type 1 correspond à une droite définie par une transformation log-linéaire :

$$\log [P(R_p \geq R)] = \alpha \log(\underline{R}) - \alpha \log(R)$$

Par suite, en notant N l'effectif total de la population, la distribution du nombre $N(R)$

de ménages ayant un *revenu inférieur* à R est définie par :

$$\begin{aligned} N(R) &= N \left[1 - \left(\frac{R}{\underline{R}} \right)^\alpha \right], \quad \forall R \geq \underline{R}, \text{ et } \alpha > 0. \\ n(R) &\equiv NF'_{R_p}(R) = \left(\frac{R^\alpha}{R^{\alpha+1}} \right) \alpha N. \end{aligned} \quad (3.48)$$

L'expression de la fonction de hasard pour la distribution du nombre $N(R)$ de ménages ayant un *revenu inférieur* à R correspond à celle donnée à l'équation (3.47).

L'entreprise trouve profitable de ne pas exclure de consommateurs. Pour que le monopole trouve profitable de ne pas exclure de consommateurs, il faut qu'il aie envie de baisser les prix p lorsque $R_0(p) > \underline{R}$. Donc, d'après la condition du premier ordre de la maximisation du profit (3.10), on doit avoir

$$\frac{\partial \Pi(p)}{\partial p} = N - N(\underline{R}) - (p - c) n(\underline{R}) \left(\frac{\partial R_0}{\partial p} \right) < 0,$$

pour tout p tel que $R_0(p) > \underline{R}$. En particulier, quand le prix p est tel que $R_0(p) = \underline{R}$ (que l'on notera $p = \underline{p}$), on doit avoir

$$\frac{\partial \Pi(p)}{\partial p} = N - N(\underline{R}) - [\underline{p} - c] n(\underline{R}) \left. \frac{\partial R_0}{\partial p} \right|_{R_0=\underline{R}} \leq 0$$

ou encore

$$[\underline{p} - c] n(\underline{R}) \left. \frac{\partial R_0}{\partial p} \right|_{R_0=\underline{R}} \geq N - N(\underline{R}). \quad (3.49)$$

En multipliant et en divisant par $R_0 p$ en $R_0 = \underline{R}$ (et donc $p = \underline{p}$) le membre de gauche de cette inégalité, on obtient :

$$\left[\frac{p - c}{p} \right]_{p=\underline{p}} \frac{\underline{R} n(\underline{R})}{R_0} \left(\frac{p}{R_0} \frac{\partial R_0}{\partial p} \right) \Big|_{R_0=\underline{R}} \geq N - N(\underline{R}).$$

En divisant maintenant les deux membres de l'inégalité résultante par $N S(\underline{R})$ et en utilisant la fonction de hazard définie en (3.47), on obtient

$$\left(\frac{p-c}{p} \right)_{p=\underline{p}} \underline{Rh}(\underline{R}) \left(\frac{p}{R_0} \frac{\partial R_0}{\partial p} \right) \Big|_{R_0=\underline{R}} \geq 1.$$

Dans le cas particulier de la distribution de Pareto, on a $\varepsilon_{D/p} = \underline{Rh}(\underline{R}) = \alpha$. Par conséquent, la condition précédente s'écrit :

$$\left(\frac{p-c}{p} \right)_{p=\underline{p}} \geq \left[\alpha \left(\frac{p}{R_0} \frac{\partial R_0}{\partial p} \right) \Big|_{R_0=\underline{R}} \right]^{-1}. \quad (3.50)$$

Dans le cas particulier des préférences telles que définies à l'équation (3.3), on sait par l'équation (3.5) que

$$\frac{u}{1+u} R_0(p) \equiv p.$$

Par suite on a

$$\left(\frac{p}{R_0} \frac{\partial R_0}{\partial p} \right) = 1$$

pour tout p et donc $\underline{p} = [u/(1+u)] \underline{R}$. Par suite, le monopole trouve profitable de ne pas exclure de consommateurs où $N(R) \sim \mathcal{L}_p(\alpha, \underline{R})$ quand

$$\frac{\underline{p}-c}{\underline{p}} \geq \frac{1}{\alpha}, \quad (3.51)$$

avec $\alpha > 1$. On doit donc avoir $\underline{p} \geq [\alpha/(\alpha-1)]c$ ou encore, d'après la définition de $R_0(p)$,

$$\underline{R} \geq \left(\frac{\alpha}{\alpha-1} \right) R_0(c). \quad (3.52)$$

Une première remarque est que cette relation entre prix unitaire p^* maximisant le profit et coût unitaire c satisfait à la loi de Van der Wijk (Cowell, 2011). Le prix p^* est donc au moins égal à une proportion $\alpha/(\alpha-1)$ du coût unitaire c de fabrication du bien. Ce qui signifie que d'autant moins élevé est α , d'autant plus inégalitaire est la répartition des revenus et d'autant plus élevé est le prix p^* . On peut aussi remarquer que le paramètre α dans cette illustration correspond à l'élasticité-prix de la demande $\varepsilon_{D/p}$. Les cher-

cheurs Ibragimov *et al.* (2018) ont prouvé, sous les hypothèses de préférences croissantes, concave avec des dérivées d'ordre trois et d'ordre quatre positives, qu'un approfondissement des inégalités rend la distribution des revenus davantage inégalitaire et réduit en valeur absolue l'élasticité prix de la demande si et seulement si les fonctions de demande individuelles et leurs dérivées sont concaves. Les caractéristiques de la distribution des revenus apparaissent une fois de plus très importantes dans la fixation du prix et donc dans la détermination de la taille de la population exclue.

Notons pour $\alpha > \frac{1}{2}$, $\mathcal{G} = (2\alpha - 1)^{-1}$ l'indice d'inégalité de Gini et $I_L = (p^* - c)/p^*$ l'indice de Lerner. L'inégalité (3.51) se ré-écrit :

$$\begin{aligned} \frac{p^* - c}{p^*} &\geq \frac{1}{\alpha} \iff \\ \frac{p^* - c}{p^* + c} &\geq \frac{1}{2\alpha - 1} \equiv \mathcal{G} \end{aligned}$$

ou de manière équivalente

$$\left(1 + \frac{c}{p^*}\right)^{-1} I_L \geq \frac{1}{2\alpha - 1} \equiv \mathcal{G} \quad (3.53)$$

En utilisant les équations (3.43) et (3.44) et sous l'hypothèse qu'il n'y a pas d'exclusion c'est-à-dire quand $p^* \leq \underline{p}$, on obtient

$$\left(\frac{\alpha}{\alpha - 1}\right) c \leq \frac{u}{1 + u} \underline{R}$$

ou encore

$$\alpha \geq \left[1 - \left(\frac{1 + u}{u}\right) \frac{c}{\underline{R}}\right]^{-1}. \quad (3.54)$$

Comme $\alpha > 0$, cette condition a un sens que si

$$1 - \left(\frac{1 + u}{u}\right) \frac{c}{\underline{R}} < 0$$

ce qui correspond à

$$\underline{R} > \left(\frac{1 + u}{u}\right) c.$$

Donc, le revenu minimum nécessaire à l'élimination de l'exclusion doit être au moins égal à un minimum de revenu représentant une fraction plus que proportionnelle du coût unitaire de production du bien. Ce qui sous-entend que ce minimum de revenu est celui qui met les ménages intéressés en situation d'acheter le bien et d'en retirer une utilité significative.

Appréciation graphique ($R_0 = \underline{R}$)

- Le prix optimal minimal maximisant le profit : $p_m = \left(\frac{\alpha}{\alpha-1}\right) c$, $\alpha > 1$.
- Le prix minimal correspondant à \underline{R} : $\underline{p} = \left(\frac{u}{1+u}\right) \underline{R}$
- Le revenu $R_0 \equiv p \left(\frac{1+u}{u}\right) = \left(\frac{\alpha}{\alpha-1}\right) \left(\frac{1+u}{u}\right) c$
- La demande : $D(\underline{p}) = N - N(\underline{R}) = N$
- Soit $\bar{E}(\underline{R})$ la fonction définie à partir du membre de gauche de l'inégalité (3.49).

Alors,

$$\bar{E}(\underline{R}) = (\underline{p} - c) n(\underline{R}) \left[\frac{v'(\underline{R} - \underline{p})}{v'(\underline{R} - \underline{p}) - v'(\underline{R})} \right].$$

Comme $v(R) = \ln(1+u) + \ln(R-p) \implies$

$$v'(\underline{R}) = \frac{1+u}{\underline{R}}, \quad v'(\underline{R} - \underline{p}) = \frac{1+u}{(1-u)\underline{R}}$$

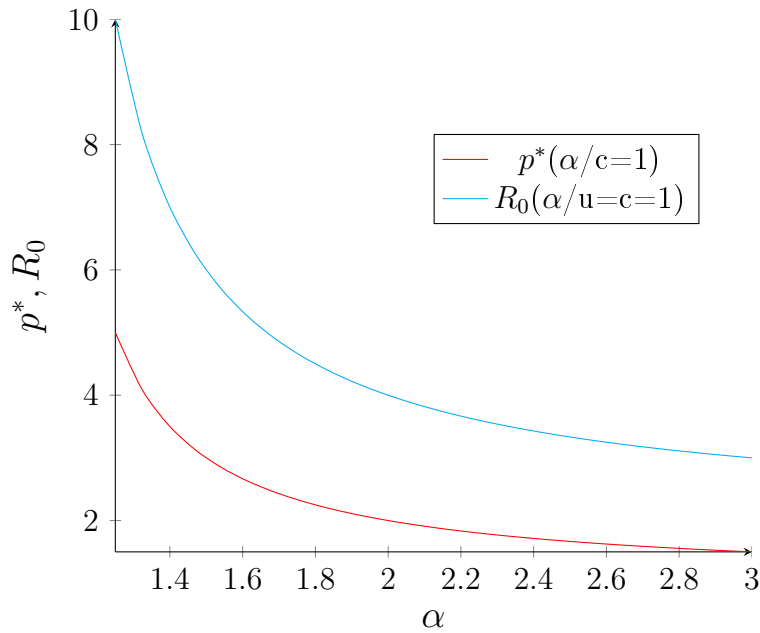
Par suite,

$$\bar{E}(\underline{R}) = \alpha N \left[\frac{(\underline{R} - c)u - c}{u\underline{R}(1+u)} \right].$$

Si on relève continuellement \underline{R} , la différence entre $\bar{E}(\underline{R})$ et $D(\underline{p})$ est évaluée à :

$$\lim_{\underline{R} \rightarrow \infty} \left\{ \alpha N \left[\frac{(\underline{R} - c)u - c}{u\underline{R}(1+u)} \right] - N \right\} = \left(\frac{\alpha - u - 1}{1+u} \right) N.$$

Cette différence sera toujours positive que si $\alpha \geq (1+u)$, $u > 0$. Une condition générale sur α pour que le monopole trouve profitable de ne pas exclure les

FIGURE 3.5.2 – Paramètre α , Prix optimal p^* et Revenu R_0 

consommateurs est telle que

$$\alpha N \left[\frac{(\underline{R} - c)u - c}{u\underline{R}(1 + u)} \right] \geq N$$

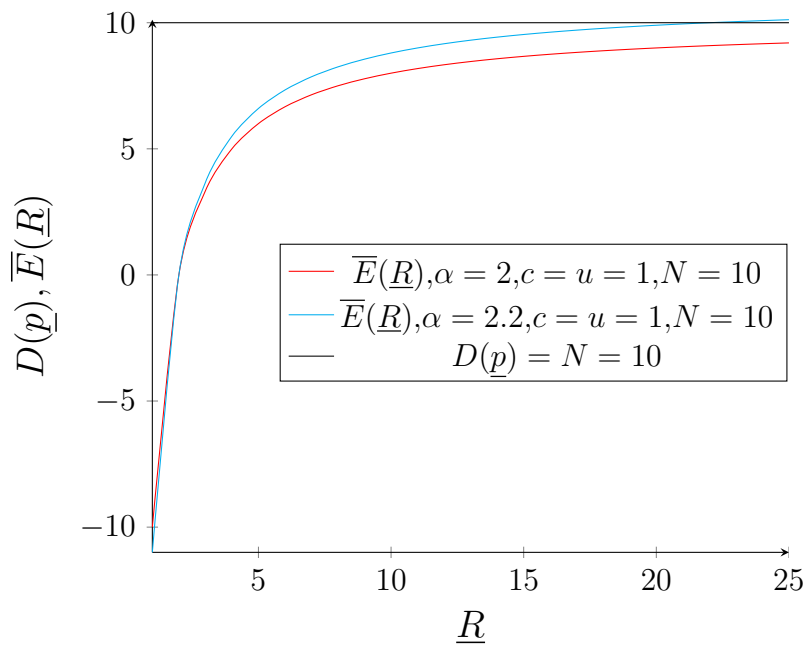
$$\alpha \geq \frac{u\underline{R}(1 + u)}{(\underline{R} - c)u - c} = \frac{\underline{R}(1 + u)}{\underline{R} - c - \frac{c}{u}} \quad (3.55)$$

à condition que $\underline{R} - c - \frac{c}{u} > 0$ c'est-à-dire

$$\underline{R} > (1 + u) \frac{c}{u}.$$

On peut donc observer que le prix optimal minimal p^* maximisant le profit de l'entreprise et le revenu minimal R_0 donnant accès au bien - R_0 restant supérieur à p^* - diminuent au fur et à mesure que α augmente ou que diminue l'inégalité des revenus (figure 3.5.2).

Une condition pour que le monopole trouve profitable de ne pas exclure les consommateurs est que \underline{R} soit suffisamment grand. Pour $\alpha = \{2, 2.2\}$, (Lubrano, 2013) le comportement monotone croissant de $\overline{E}(\underline{R})$ (figure 3.5.3) en fonction de \underline{R} par rapport à $D(\underline{p})$ confirme

FIGURE 3.5.3 – $\bar{E}(\underline{R})$ et $D(\underline{p})$ en fonction de \underline{R} 

cette idée selon laquelle un relèvement soutenu de \underline{R} contribue progressivement à l'élimination de l'exclusion ($\bar{E}(\underline{R}) \rightarrow D(\underline{p})$).

3.5.3.2 Distribution log-normale

La distribution log normale est mieux adaptée pour modéliser les revenus faibles et les revenus moyens et possède de propriétés intéressantes pour l'analyse de la pauvreté (Lubrano, 2013; Cowell, 2011; Kalecki, 1945).

Encadré 3.5.3.2 : Distribution lognormale.

La distribution lognormale possède les propriétés suivantes :

- La distribution des revenus soumis à un processus aléatoire multiplicatif, converge vers une distribution log normale
- La variable $\log(R)$ peut être assimilée à un indice de bien-être. (*Gilbrat, Van Praag, Herwaarden, Kapteyn*)
- Les mesures d'inégalité tels les indices de Gini et de Pietra ^a ne dépendent que du paramètre mesurant la dispersion σ .
- La distribution log normale s'ajuste raisonnablement bien à des séries de données de revenus.

^a. Dans le cas d'une distribution log-normale, l'indice de Pietra s'écrit : $P_R = [2\Phi(\sigma^2/2) - 1]$ et l'indice de Gini : $G = [2\Phi(\sigma/\sqrt{2}) - 1]$. La fonction $\Phi(\cdot)$ correspond à fonction de répartition d'une distribution normale $\mathcal{N}(0, 1)$.

1. Définition. Soit X une variable aléatoire $\mathcal{N}(0, 1)$ telle que ses fonctions de densité $\varphi(x)$ et de répartition $\Phi(x)$ sont respectivement définies par :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}, \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds.$$

Soit R le montant de revenu détenu par un individu quelconque d'une population donnée. La variable aléatoire $R_{\mathcal{L}}$ suit une loi log normale de paramètres (μ, σ^2) si $\ln(R_{\mathcal{L}}) \sim \mathcal{N}(\mu, \sigma^2)$. Sa fonction de densité $f_{R_{\mathcal{L}}}(R)$ et sa fonction de répartition $F_{R_{\mathcal{L}}}(R)$ sont respectivement définies par :

$$f_{R_{\mathcal{L}}}(R) \equiv \varphi\left(\frac{\ln R - \mu}{\sigma}\right) = \frac{1}{\sigma R \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln R - \mu}{\sigma}\right)^2}, R \in \mathbb{R}_+^*.$$

$$F_{R_{\mathcal{L}}}(R) \equiv \int_{-\infty}^R f_{R_{\mathcal{L}}}(S) dS = \Phi\left(\frac{\ln R - \mu}{\sigma}\right)$$

telles que

$$\mu_{\mathcal{L}} \equiv E(R_{\mathcal{L}}) = e^{\mu + \frac{\sigma^2}{2}}, \sigma_{\mathcal{L}}^2 \equiv V(R_{\mathcal{L}}) = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}.$$

On note : $R_{\mathcal{L}} \sim \mathcal{LN}(\mu, \sigma^2)$.

2. Les fonctions de survie et de hasard sont respectivement définies par :

$$\begin{aligned} S(R) &\equiv 1 - F_{R_{\mathcal{L}}}(R) = 1 - \Phi\left(\frac{\ln R - \mu}{\sigma}\right), \\ h(R) &\equiv \frac{f_{R_{\mathcal{L}}}(R)}{S(R)} = \frac{1}{R\sigma} \left[\frac{\varphi\left(\frac{\ln R - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\ln R - \mu}{\sigma}\right)} \right] \end{aligned} \quad (3.56)$$

de sorte que

$$Rh(R) = \frac{1}{\sigma} \left[\frac{\varphi\left(\frac{\ln R - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\ln R - \mu}{\sigma}\right)} \right].$$

Par conséquent, la fonction de répartition du nombre $N(R)$ de personnes ayant un revenu inférieur à R et la fonction de densité $n(R)$ associée sont respectivement définies par :

$$\begin{aligned} N(R) &= N\Phi\left(\frac{\ln R - \mu}{\sigma}\right), \quad R \in \mathbb{R}_+^*, \\ n(R) &= \frac{N}{\sigma R} \varphi\left(\frac{\ln R - \mu}{\sigma}\right). \end{aligned} \quad (3.57)$$

L'expression de la fonction de hasard pour la distribution du nombre de personnes ayant un revenu inférieur à R , correspond à celle donnée à l'équation (3.56). Dans ce cas, pour que le monopole trouve profitable de ne pas exclure de consommateurs, on doit avoir :

$$\left(\frac{p^* - c}{p^*}\right) \left\{ \frac{1}{\sigma} \left[\frac{\varphi\left(\frac{\ln R - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\ln R - \mu}{\sigma}\right)} \right] \right\} \left(\frac{p}{R_0} \frac{\partial R_0}{\partial p} \right) \Big|_{R_0 = \underline{R}} \geq 1 \quad (3.58)$$

ou de manière équivalente¹⁸

$$\begin{aligned} \frac{p^* - c}{p^*} &\geq \frac{1 - \Phi\left(\frac{\ln R - \mu}{\sigma}\right)}{\varphi\left(\frac{\ln R - \mu}{\sigma}\right)} \sigma. \\ &\geq \lambda(R_c) \sigma, \quad \lambda(R_c) = \frac{1 - \Phi(R_c)}{\varphi(R_c)}, \quad R_c = \frac{\ln R - \mu}{\sigma}. \end{aligned} \quad (3.59)$$

Étant donné que la variable $(\ln R_{\mathcal{L}}) \sim \mathcal{N}(\mu, \sigma^2)$ alors

$$\begin{aligned} E[\ln R_{\mathcal{L}} / (\ln R_{\mathcal{L}} > R_c)] &= \mu + \sigma \frac{\varphi(R_c)}{1 - \Phi(R_c)} \\ \sigma^2 \{E[\ln R_{\mathcal{L}} / (\ln R_{\mathcal{L}} > R_c)] - \mu\}^{-1} &= \frac{1 - \Phi(R_c)}{\varphi(R_c)} \sigma. \end{aligned}$$

Par suite,

$$\begin{aligned} \frac{p^* - c}{p^*} &\geq \lambda(R_c) \sigma, \quad R_c = \frac{\ln R - \mu}{\sigma} \\ &\geq \sigma^2 \{E[\ln R_{\mathcal{L}} / (\ln R_{\mathcal{L}} > R_c)] - \mu\}^{-1}. \end{aligned}$$

Si on note $\mu_c = E[\ln R_{\mathcal{L}} / (\ln R_{\mathcal{L}} > R_c)]$ c'est-à-dire la moyenne du logarithme des revenus supérieurs à R_c et $\Delta\mu = \mu_c - \mu$, l'écart entre la moyenne des revenus supérieurs et la moyenne globale des revenus, la condition de non-exclusion s'écrit :

$$\frac{p^* - c}{p^*} \geq \frac{\sigma^2}{\Delta\mu}. \quad (3.60)$$

Dès lors, en supposant $\Delta\mu$ fixé - les paramètres μ_c et μ variant dans une même proportion - l'ampleur de l'inégalité s'apprécie à travers le comportement du paramètre σ^2 mesurant

18. Le fonction $\lambda(R_c)$ correspond au *Ratio de Mills*. Ce ratio qui a pour expression (Gasull et Utzet, 2014)

$$\begin{aligned} \lambda(R_c) &= \frac{1 - \Phi(R_c)}{\varphi(R_c)} \\ &= e^{\frac{R_c^2}{2}} \int_{R_c}^{+\infty} e^{-\frac{M^2}{2}} dM, \end{aligned}$$

vérifie les propriétés suivantes : $\lambda(R_c) \geq 0$, $\lambda'(R_c) > 0$, $\lim_{R_c \rightarrow -\infty} \lambda(R_c) = 0$ et $\lim_{R_c \rightarrow \infty} \lambda'(R_c) = 1$.

la dispersion des revenus. D'autant plus dispersés sont les revenus, d'autant plus élevé sera l'écart relatif minimal du prix optimal et d'autant plus important sera le nombre de ménages exclus de la consommation du bien. A l'inverse, d'autant moins dispersés sont les revenus, d'autant moins élevé sera l'écart relatif minimal du prix optimal et d'autant moins important sera le nombre de ménages exclus de la consommation du bien. On peut encore mieux apprécier l'effet du degré de l'inégalité mesuré par l'indice \mathcal{G} de Gini sur le nombre de ménages exclus en réécrivant l'inégalité (3.60).

On sait que (Lubrano, 2013) pour une distribution log-normale,

$$\mathcal{G} = 2\Phi\left(\sigma/\sqrt{2}\right) - 1$$

et donc, $\Phi^{-1}(u)$ étant la quantile d'ordre u de la loi normale centrée réduite,

$$\sigma = \sqrt{2}\Phi^{-1}\left(\frac{\mathcal{G} + 1}{2}\right), \Phi^{-1}\left(\frac{\mathcal{G} + 1}{2}\right) \in]\Phi^{-1}(1/2); \Phi^{-1}(1)[$$

où $(\frac{\mathcal{G}+1}{2}) \in]1/2; 1[$ étant donné que $G \in]0; 1[$. Avec une précision de 10^{-4} ,

$$]\Phi^{-1}(1/2); \Phi^{-1}(1)[\longrightarrow]0; 4[.$$

Par suite, l'inégalité (3.60) se ré-écrit

$$\frac{p^* - c}{p^*} \geq \frac{2[\Phi^{-1}(\frac{\mathcal{G}+1}{2})]^2}{\Delta\mu}. \quad (3.61)$$

Donc, pour $\Delta\mu$ fixé, d'autant plus inégalitaire est la distribution des revenus ($\mathcal{G} \rightarrow 1$), d'autant plus élevé est le rapport $\left\{2[\Phi^{-1}(\frac{\mathcal{G}+1}{2})]^2/\Delta\mu\right\}$, d'autant plus élevée est la marge perçue par l'entreprise. Ce qui correspond à l'analyse faite à partir de l'inégalité (3.53) obtenue dans le cas de la distribution de Pareto.

3.5.4 Généralisation des résultats du modèle

Les illustrations numériques effectuées ont permis de comprendre que la production du bien X peut se réaliser dans un contexte de moindre exclusion possible. En effet, le monopole peut produire de manière viable le bien X à un prix unitaire $p^* \geq c$, c étant le coût unitaire, accessible aux consommateurs intéressés ayant un revenu minimum $R_0(p)$. L'importance du niveau d'exclusion dépend de l'écart entre $R_0(p)$ et le revenu minimum \underline{R} de la société, écart qui s'ajuste, comme l'étaient les illustrations effectuées, en fonction des caractéristiques de la distribution des revenus au niveau de cette société. A ce niveau, il s'agit d'identifier formellement les conditions sur la distribution des revenus $n(R)$ qui sont telles que les prix définis en (3.11) et (3.44) vérifient $p^* < \underline{p}$.

3.5.4.1 Demande unitaire et revenu

Définition et caractérisation d'une fonction d'utilité indirecte. Dans ce paragraphe, on se propose de définir et de caractériser formellement une fonction d'utilité indirecte générale qui soit cohérente avec les préférences adoptées en (3.4.1).

On considère une fonction d'utilité indirecte $\mathcal{V}(p, R)$ définie par :

$$\mathcal{V}(p, R) = \begin{cases} u + v(R - p) & \text{si on achète le bien} \\ v(R) & \text{si on renonce à son achat} \end{cases} \quad (3.62)$$

où $u > 0$ est l'utilité retirée de consommation d'une unité du bien X , p correspondant au prix unitaire du bien et R le revenu des ménages. On suppose que :

Hypothèse 3.5. *La fonction $v(\cdot)$ est telle que $v'(\cdot) > 0$ et $v''(\cdot) < 0$.*

La fonction d'utilité indirecte $v(\cdot)$ telle que définie en (3.62) est donc supposée deux fois dérivable, croissante et concave.

Un revenu minimum est indispensable à la survie alors que le bien (unitaire) considéré ne l'est pas. De ces considérations, on suppose raisonnablement que :

Hypothèse 3.6. *Le bien R est un bien essentiel et X ne l'est pas et donc*

$$\begin{aligned} \lim_{R \rightarrow 0} v(R) &= -\infty, \\ \lim_{R \rightarrow +\infty} v'(R) &= 0, \quad (\text{utilité marginale du revenu décroissante}) \\ \lim_{x \rightarrow 0^+} \frac{\partial \mathcal{V}}{\partial x} &= \bar{u} < +\infty. \end{aligned}$$

On peut en déduire de ces hypothèses, que le ménage ayant un niveau de revenu qui lui permet d'acheter le bien, en retire une utilité de plus en plus importante que celui qui ne l'achète pas à mesure que le revenu augmente.

Proposition 3.7. *Pour une fonction d'utilité indirecte telle que définie à l'équation (3.62) associée à $\mathcal{U}(X, R)$ et satisfaisant les hypothèses (3.1) et (3.3), il existe un revenu critique $R_0(p, u) > p$, tel que tous les ménages ayant un revenu supérieur (inférieur) à ce revenu critique auront (n'auront pas) accès au bien X . De plus, ce revenu seuil augmente plus que proportionnellement avec p et décroît avec u . Formellement :*

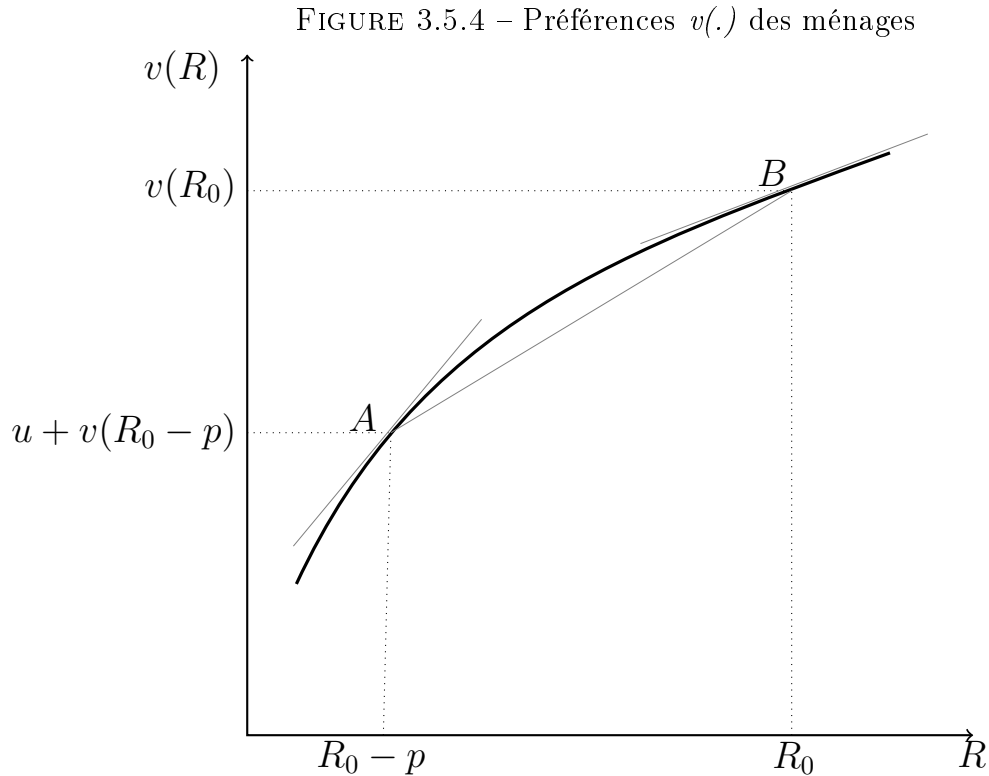
$$\frac{dR_0}{dp} > 1 \quad \text{and} \quad \frac{dR_0}{du} < 0.$$

Démonstration. Annexe 3.A.4, 3.A.4 □

3.5.4.2 Caractérisation d'une fonction $R_0(p)$, d'une fonction $D(p)$ et conditions d'accès au bien à tous

La détermination d'un revenu minimum $R_0(p)$ et d'une fonction de demande $D(p)$ constituent, dans le cadre d'une distribution de revenus donnée, deux fonctions importantes pour définir et caractériser les conditions d'accès au bien des ménages à faible revenus.

On considère une population de taille N où les ménages ayant des préférences similaires diffèrent quant à leurs revenus. On suppose que les revenus sont distribués dans l'intervalle



$[\underline{R}; \bar{R}]$, $\bar{R} (\rightarrow \infty)$ étant le revenu maximum de la société. Soit $m = \underline{N}/N$ la fraction de la population ayant le revenu \underline{R} . Soient $G(R)$ et $N(R) = NG(R)$ définissant respectivement la fonction de densité cumulative des revenus et le nombre de ménages ayant au plus un revenu R .

Les quatre propositions suivantes définiront et caractériseront $R_0(p)$, $D(p)$ et les conditions d'accès au bien des ménages à faible revenus.

Proposition 3.8. *On suppose que $R_0 \in [\underline{R}; \bar{R}]$. Le revenu $R_0(p)$ est une fonction croissante de p et la demande agrégée $D(p)$ est décroissante en p et définie par :*

$$D(p) \equiv \int_{R_0}^{+\infty} n(R) dR = N - N(R_0), \text{ pour tout } p \geq \underline{p}.$$

$$= \begin{cases} N(1-m)[1-G(R_0)] & \text{pour tout } p \geq \underline{p}. \\ N & \text{sinon.} \end{cases}$$

où $n(R) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est la fonction de densité du nombre d'individus ayant un revenu R et $n(R) = N(1-m)g(R)$, $g(R)$ étant la fonction de densité associée à $G(R)$.

Démonstration. Annexe 3.A.4 et proposition 3.8 □

Dans le cadre de ce modèle de demande unitaire, on peut faire remarquer que l'élasticité-prix de la demande se définit en relation avec la distribution du revenu des ménages. De manière précise, pour tout $p \geq \underline{p}$, l'élasticité-prix de la demande est donnée par : (Annexe : équation 3.77)

$$\epsilon_{D/p} = \left(\frac{pR'_0(p)}{R_0(p)} \right) \left(\frac{R_0(p)g(R_0)}{1-G(R_0)} \right) \quad (3.63)$$

Cette élasticité s'exprime donc comme le produit de deux élasticités : celle donnant le revenu minimum $R_0(p)$, qu'il faut à un ménage pour disposer à acheter le bien par rapport au prix et celle correspondant au nombre de consommateurs (ménages) $N(1-G(R_0))$ par rapport à $R_0(p)$ ayant accès au bien. On peut remarquer que $g(R_0)/(1-G(R_0))$ correspond exactement à la fonction de hasard donnant la proportion de ménages percevant un revenu R_0 parmi ceux ayant au plus ce revenu.

Proposition 3.9. *On note c le coût unitaire de fabrication du bien X . Les profits du monopole s'écrivent donc :*

$$\Pi(p) = (p - c) D(p, R).$$

Soit $F(p, R)$ une fonction continue et dérivable sur $\mathbb{R} \times \mathbb{R}_+^n$ définie par :

$$F(p, R) \equiv \frac{\partial \Pi(p)}{\partial p} = D(p, R) + (p - c) \frac{\partial D(p, R)}{\partial p}$$

et soit $p^*(\mathbf{R}) : \mathbb{R}_+^n \rightarrow \mathbb{R}$ la fonction implicite définie par la solution en p de l'équation $F(p, R) = 0$. Si $F(p, \cdot)$ est Schur-convexe (resp. concave)¹⁹ pour tout $p \in \mathbb{R}$ et $[\partial F(p, R)/\partial p] < 0$ pour tout $\mathbf{R} \in \mathbb{R}_+^n$ alors la fonction $p^*(\mathbf{R})$ est Schur-concave (resp. convexe).

Démonstration. Annexe 3.A.4, proposition 3.9.

19. Par définition, la fonction $F(p, \cdot)$ définie par

$$\mathbf{R} \in \mathbb{R}_+^n \mapsto F(p, \mathbf{R}) \in \mathbb{R}$$

est Schur-convexe (resp. concave) quand, pour tout paramètre donné $p \in \mathcal{P}$

$$\mathbf{R} \prec \mathbf{R}' \Rightarrow F(p, \mathbf{R}) \leq (\geq) F(p, \mathbf{R}').$$

Cette proposition établit qu'au prix p fixé, la fonction $F(p, R)$ étant schur-convexe, croit avec les inégalités. A l'inverse, si les inégalités augmentent, le prix optimal $p^*(R)$ obtenu quand $F(p, R) \equiv 0$, augmente aussi. \square

Proposition 3.10. *On note c le coût unitaire de fabrication du bien X . Les profits du monopole s'écrivent donc :*

$$\Pi(p) = (p - c) [N - N(R_0)]. \quad (3.64)$$

Si $R_0(p) \leq \underline{R}$, lorsqu'un monopole qui maximise ses profits adopte un prix qui n'entraîne pas d'exclusion, alors $p = \underline{p}$ où \underline{p} est défini par l'équation $R_0(\underline{p}) = \underline{R}$.

Démonstration. Annexe 3.A.4 et proposition 3.10

Donc, tant qu'il n'y a pas exclusion, le monopole a intérêt à augmenter le prix p de sorte que R_0 va augmenter jusqu'à atteindre \underline{R} . \square

Proposition 3.11. *Le prix unique p vérifiant $R_0(p) = \underline{R}$ est déterminé dans le cadre d'une distribution des revenus telle que le monopole trouve profitable de ne pas exclure de consommateurs .*

Démonstration. Proposition 3.11, Annexe 3.A.4

Dans le cadre d'une distribution des revenus telle que le monopole trouve profitable de ne pas exclure de consommateurs, une condition générale garantissant un accès au bien à tous s'écrit :

$$h(R_0) \equiv \frac{n(R_0)}{N - N(R_0)} \geq \frac{v'(R_0 - p) - v'(R_0)}{(p - c)v'(R_0 - p)} \equiv w(R_0) \quad (3.65)$$

telle que si $R_0(p) = \underline{R}$ alors $N(R_0) = 0$. \square

Proposition 3.12. *Sous l'hypothèse additionnelle que $v'''(.) > 0$ ²⁰, la fonction*

$$w(R_0) = \left(\frac{1}{p-c} \right) \left[\frac{v'(R_0 - p) - v'(R_0)}{v'(R_0 - p)} \right], \quad R_0(p) > p > c$$

est monotone décroissante.

Démonstration. Proposition 3.A.4 : Annexes 3.A.4 et 3.A.4 □

Les fonctions $h(R_0)$ et $w(R_0)$ sont donc deux fonctions décroissantes²¹. On s'attend à ce qu'à mesure que R augmente, $h(R_0)$ et $w(R_0)$ diminuent progressivement jusqu'à ce que, pour tout R_0 , $h(R_0) \geq w(R_0)$ traduisant ainsi l'élimination de l'exclusion.

Illustration. Soit $n(R) \sim L_P(\alpha, R)$ et $v(R) = u + \ln(R - p)$, $R > p > c$. Dans ce cas,

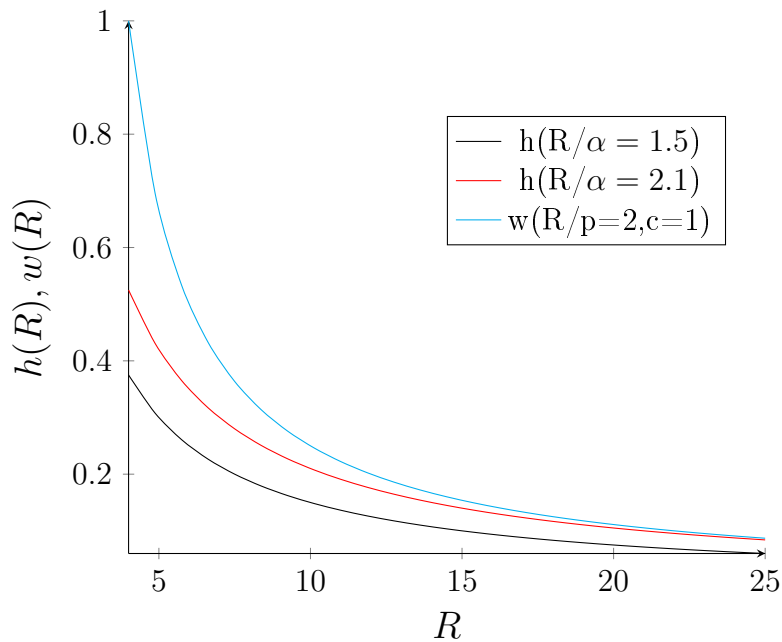
$$\begin{aligned} w(R_0) &= \left(\frac{1}{p-c} \right) \frac{v'(R_0 - p) - v'(R_0)}{v'(R_0 - p)} \\ &= \left(\frac{1}{p-c} \right) \left[\frac{1/(R_0 - p) - 1/(R_0)}{1/(R_0 - p)} \right] \\ &= \left(\frac{p}{p-c} \right) \frac{1}{R_0} \end{aligned}$$

et donc

$$h(R) = \frac{\alpha}{R} \text{ et } w(R) = \frac{p}{(p-c)R}.$$

20. Dans la littérature traitant de l'analyse du risque, la dérivée d'ordre trois est utilisé pour caractériser le comportement des individus dits « averses au risque de pertes potentielles » (downside risk averse). Selon un article publié par Menezes *et al.* (1980), ce type d'aversion au risque décrit le comportement attribué à un motif d'évitement d'une catastrophe. Ces individus sont alors ceux qui évitent des situations risquées offrant de potentiels gains substantiels du fait que ces situations rendent ces individus vulnérables à des pertes en dessous d'un certain seuil. Eeckhoudt et Gollier (2005); Kimball (1990) qualifient ces individus averses au risque de pertes potentielles de « prudents » c'est-à-dire des individus qui, s'étant obligés à faire face à un risque accru, préfèrent lier ce risque aux meilleurs résultats qu'aux pires. La caractérisation du comportement « prudent » est établi à travers le corollaire 1 stipulant qu'un individu est prudent si et seulement si son utilité espérée vérifie $u'''(.) > 0$. Différentes recherches tant théoriques qu'empiriques utilisent ce résultat caractérisant le comportement des individus prudents (Deck et Schlesinger, 2018; Crainich *et al.*, 2015).

21. La fonction $h(R)$ peut être décroissante, constante ou croissante. Elle vérifie les propriétés suivantes : $h(R) \geq 0$ et $\int_0^{+\infty} h(R) dR \rightarrow +\infty$.

FIGURE 3.5.5 – Comportement de $h(R)$ et $w(R)$ 

En particulier, quand $h(R) = w(R)$, la comparaison de $h(R)$ et de $w(R)$ ne dépend pas de R et l'on a :

$$\frac{p - c}{p} = \frac{1}{\alpha}.$$

Le tracé de ces deux fonctions (figure 3.5.5) étaye bien l'idée selon laquelle un relèvement continu du revenu \underline{R} conduit progressivement à l'élimination de l'exclusion. Pour de faibles valeurs de R , on $h(R) < w(R)$. A mesure que R augmente, $h(R)$ et $w(R)$ diminuent progressivement jusqu'à ce que, pour tout R , $h(R) \geq w(R)$. Ce processus d'élimination de l'exclusion converge d'autant plus vite que l'inégalité des revenus est moins importante.

Réécriture de la condition définie à l'équation (3.84). Par la proposition (3.10), il est établi que, quand *il n'y a pas d'exclusion*, le monopole qui maximise son profit a intérêt à adopter un prix p tel que $R_0 = \underline{R}$. Dans ce cas de figure, la condition (3.65) s'écrit :

$$\frac{n(\underline{R})}{N - N(\underline{R})} \geq \frac{v'(\underline{R} - p) - v'(\underline{R})}{(p - c)v'(\underline{R} - p)}$$

où \underline{R} , et en conséquence \underline{p} , est tel que $[u + v(\underline{R} - \underline{p})] = v(\underline{R})$. En explicitant \underline{p} de cette égalité, il vient : $\underline{p} = \underline{R} - v^{-1}[v(\underline{R}) - u]$, $v^{-1}(\cdot)$ étant l'image inverse de $v(\cdot)$. Par suite,

$$\frac{n(\underline{R})}{N - N(\underline{R})} \geq \frac{v'(\underline{R} - \underline{p}) - v'(\underline{R})}{\{\underline{R} - c - v^{-1}[v(\underline{R}) - u]\} v'(\underline{R} - \underline{p})}. \quad (3.66)$$

La proposition 3.11 établit sans équivoque que, dans un monde sans exclusion, le prix optimal p du bien s'ajuste en fonction de l'élasticité-prix de la demande des ménages ayant le revenu minimum \underline{R} . De manière précise, l'écart relatif de prix s'ajuste en fonction de la sensibilité des consommateurs et en particulier des ménages à faibles revenus. Dans un second temps, on a établi que le revenu \underline{R} satisfait au résultat dérivé en (3.7) stipulant que le revenu augmente plus que proportionnellement à l'accroissement du prix p . Enfin, nous avons pu établir une condition générale de non-exclusion liant la distribution des revenus aux préférences des ménages. Ce résultat prouve bien que les caractéristiques de la distribution des revenus constituent un facteur de première importance qui intervient dans la façon dont s'expriment les préférences des ménages. On comprend donc que, si une entreprise fait face à une population où la distribution des revenus est très inégalitaire, il lui sera au contraire profitable d'exclure certains consommateurs potentiels pour pouvoir s'assurer une marge plus importante sur les autres. Et donc, une modification adéquate de la structure de la distribution des revenus peut se révéler efficace pour donner à tous ceux ayant un revenu minimum et qui le désirent, l'accès au bien.

3.6 Les résultats du modèle eu égard à la littérature

Les principal résultat obtenu est donc l'établissement d'une condition générale de non-exclusion à partir de la définition et de la preuve de l'existence d'un niveau de revenu minimal $R_0(p)$ qui donne accès à l'achat du bien. Cette condition générale est établie après qu'il ait été démontré, pour un revenu minimum \underline{R} de la société tel que $R_0(p) \leq \underline{R}$, que l'entreprise a intérêt à produire et maximiser son profit en $R_0(p) = \underline{R}$.

Notre modèle a quelques points communs avec celui de Malueg et Schwartz (1994). L'objectif de leur modèle, il faut le rappeler, a été d'évaluer sur le bien-être l'impact

d'un régime de prix uniforme ou d'un système de prix basé sur une discrimination du troisième degré dans une situation d'importations parallèles entre pays (consommateurs) à niveaux de revenus hétérogènes. Développé dans le cadre d'un continuum d'individus, notre modèle s'est basé, à la suite de ces chercheurs, sur une fonction d'utilité séparable par rapport aux revenus supposés hétérogènes. Nous avons envisagé une distribution de revenu générale ne donnant pas forcément lieu à des demandes linéaires contrairement à ces auteurs qui présument que les revenus sont uniformément distribués. Ce qui donne lieu à des demandes linéaires. Nos premiers résultats concordent avec entre autres, les premières trouvailles de ces chercheurs. En effet, ils ont prouvé que d'autant plus dispersée est la demande - dispersion due à des différences de revenus -, d'autant plus important est le nombre de consommateurs (pays) non approvisionnés. Une différence importante avec notre modèle est que dans leur article, il n'a pas été question de déterminer les conditions dans lesquelles une firme trouve profitable de ne pas exclure les consommateurs à l'accès du bien.

Le modèle que nous développons s'inscrit aussi en partie dans la démarche suivie par He et Sun (2006) dans leur publication. Ces derniers ont aussi utilisé une fonction d'utilité séparable par rapport au revenu et analysé dans un premier temps comment l'hétérogénéité des revenus contribue à exclure les individus à faibles revenus de la consommation du bien. Notre modèle a deux différences importantes. Contrairement au modèle discret de He et Sun (2006), nous avons considéré un continuum de consommateurs. De plus, ces chercheurs, à la suite de Malueg et Schwartz (1994), ont supposé que les revenus sont uniformément distribués ce qui donne lieu à des demandes linéaires. Comme signalé précédemment, nous avons considéré une distribution de revenus générale. Ce qui nous a permis de dériver une fonction de demande agrégée plus générale. Bien que la méthodologie adoptée et les objectifs poursuivis sont différents, nos résultats concordent avec ceux de He et Sun (2006) , entre autres résultats, sur l'impact de la dispersion et de l'inégalité des revenus sur le nombre d'individus exclus. En effet, sous un régime de prix uniforme, ils ont prouvé que d'autant plus élevées sont la dispersion et l'inégalité des revenus, d'autant plus important est le nombre d'individus exclus de la consommation du bien. Notre

recherche se démarque aussi de celle de ces chercheurs par la dérivation d'une inégalité liant la distribution des revenus des ménages à leurs préférences.

La fonction d'utilité utilisée dans notre cas s'apparente à celle adoptée par Gabszewicz et Thisse (1979) qui analysent l'effet de la dispersion de revenus sur la formation différenciée des prix, l'offre de produits différenciés et sur le nombre de marchés approvisionnés dans le cadre d'un marché duopolistique. Ces chercheurs, bien que leur but a été totalement différent du nôtre, ont souligné que les différences significatives de revenus entre les consommateurs jouent un rôle crucial dans l'offre de produits différenciés. Ce résultat démontré par Gabszewicz et Thisse (1979) a été généralisé par Acharyya (2005) qui démontre que les différences significatives de revenus des consommateurs impliquant des contraintes différentes de budget, constituent une variable importante contribuant à une discrimination par les entreprises basée sur la qualité. Évidemment, le but n'a pas été non plus d'identifier les conditions dans lesquelles une entreprise trouve profitable de ne pas exclure des consommateurs à faibles revenus.

Notre contribution essentielle dans cet essai réside dans l'élaboration et la caractérisation des circonstances dans lesquelles une firme trouve profitable de ne pas exclure des consommateurs (propositions 3.11, 3.12, 3.9, 3.11) partant de la formulation d'une fonction d'utilité générale qui n'est pas quasi-linéaire.

3.7 Synthèse des principaux résultats

L'objectif essentiel de cet essai a été d'identifier les conditions dans lesquelles une entreprise productrice de biens et services trouve profitable de ne pas exclure de consommateurs. Pour arriver à cette fin, nous avons adopté une fonction d'utilité qui n'est pas quasi-linéaire prenant en compte explicitement l'effet du revenu des individus sur leurs préférences. La demande agrégée a été déterminée une fois définis le niveau de revenu minimal R_0 donnant accès au bien et la distribution du nombre d'individus $n(R)$ ayant un niveau de revenu R . En supposant que le coût de production c d'une unité du bien est tel que $p \geq c$, p étant le prix unitaire, le profit de l'entreprise et le bien-être social (marshallien) sont ensuite calculés.

Le prix de vente optimal p^* qui maximise le profit vérifie l'égalité classique de Lerner $[(p^* - c) / p^*] = 1 / \varepsilon_{D/p}$ signifiant ainsi que ce prix sera d'autant plus élevé que la demande est inélastique. La disponibilité de deux qualités du bien (monopole à produits différenciés ou duopole constitué de deux firmes compétitives) conforte ce résultat trouvé dans le cas du monopole. Ce prix qui maximise le profit est évidemment trop élevé par rapport au prix qui maximise le bien-être social. Il faut souligner que la demande dont il s'agit est celle des individus ayant niveau de revenu minimal R_0 . Ce qui traduit que les individus ayant un niveau de revenu $R < R_0$ sont exclus de la consommation du bien. Dans l'optique d'une maximisation du bien-être social, le nombre d'exclus sera évidemment moins important. Néanmoins, le prix maximisant le bien-être social ne garantit pas forcément un accès au bien à tous les ménages ayant un revenu $R < R_0$ qui le souhaitent.

Les résultats théoriques trouvés ont été illustrés en supposant d'abord que $n(R)$ est Pareto-distribué et ensuite Log normale-distribué. Dans le cas de la distribution de Pareto²², il est révélé que le prix optimal minimal chargé par l'entreprise est d'autant plus élevé que les revenus sont inégalement répartis au niveau de la société. Vu sous l'angle de la dispersion des revenus, la distribution log-normale a permis de constater que le prix optimal minimal chargé par l'entreprise est d'autant plus élevé que les revenus sont dispersés.

A partir de la formulation d'une fonction d'utilité (indirecte) qui n'est pas quasi-linéaire et à caractère général, les résultats sont généralisés. Nous avons d'abord prouvé que $R_0(p)$ et $D(p)$ sont respectivement une fonction croissante et décroissante de p (proposition 3.8) et que le monopole qui maximise son profit a intérêt à adopter un prix p tel que $R_0(p) = \underline{R}$ (proposition 3.10). En dépit du fait que plusieurs aspects du modèle semblent être discrets (il n'est possible d'acquérir qu'une unité du bien), la généralisation des résultats (sous-section 3.5.4) a permis d'expliciter des conditions générales d'accès aux biens et services des ménages pauvres dans le cadre d'une distribution de revenus donnée. En effet, nous avons pu démontrer que la fonction $F(p, R)$ (proposition 3.9) définie à partir des conditions du premier ordre de l'équation de profit est Schur-convexe,

22. La distribution de Pareto, de par ses propriétés, est utilisée ici du fait de la présence en Haïti de beaucoup d'individus très pauvres (60% environ), plutôt que très riches (voir encadré 3.1).

ce qui donne la garantie que cette fonction, pour p donné, croit avec les inégalités. Par la proposition 3.11, nous avons pu établir dans quelle condition il peut s'avérer profitable pour un monopole de ne pas exclure des consommateurs de la consommation du bien. Ce résultat important signifie que la problématique de non-exclusion - ou d'exclusion - est tout simplement liée à la façon dont les revenus sont distribués et répartis dans l'économie.

Bibliographie

- ACHARYYA, R. (2005). Quality discrimination among income constrained consumers. *Economics Letters*, 86(2):245–251.
- ALESINA, A. et GIULIANO, P. (2011). Preferences for redistribution. *In Handbook of social economics*, volume 1, pages 93–131. Elsevier.
- ARNOLD, B. C. (2014). Pareto distribution. *Wiley StatsRef : Statistics Reference Online*, pages 1–10.
- BAUJARD, A. (2013a). Utilitarianism and anti-utilitarianism. *GATE Working Paper*.
- BAUJARD, A. (2013b). Welfare economics. *HAL, halshs-00906907*.
- BLANK, R. M. (2003). Selecting among anti-poverty policies : can an economist be both critical and caring? *Review of Social Economy*, 61(4):447–469.
- BOZZO-REY, M. et DARDENNE, E. (2012). *Deux siècles d'utilitarisme*. Presses universitaires de Rennes.
- BRADSHAW, T. K. (2007). Theories of poverty and anti-poverty programs in community development. *Community Development*, 38(1):7–25.
- BRAOUEZEC, Y. (2012). Customer-class pricing, parallel trade and the optimal number of market segments. *International Journal of Industrial Organization*, 30(6):605–614.
- CLARK, C. (1973). The marginal utility of income. *Oxford Economic Papers*, 25(2):145–159.
- COWELL, F. (2011). *Measuring inequality*. Oxford University Press.
- CRAINICH, D., ECKHOUDT, L. R. et HAMMITT, J. K. (2015). The value of risk reduction : new tools for an old problem. *Theory and Decision*, 79(3):403–413.
- D'ASPREMONT, C. et GEVERS, L. (1977). Equity and the informational basis of collective choice. *The Review of Economic Studies*, 44(2):199–209.

- DECK, C. et SCHLESINGER, H. (2018). On the robustness of higher order risk preferences. *Journal of Risk and Insurance*, 85(2):313–333.
- DESAI, M. (2008). Profit and profit theory. *The New Palgrave Dictionary of Economics : Volume 1–8*, pages 5186–5196.
- DEUTSCH, J., SILBER, J. et VERME, P. (2012). On social exclusion in macedonia : Measurement and determinants. *Poverty and Exclusion in the Western Balkans. New Directions in Measurement and Policy. Springer*.
- ECKHOUDT, L. et GOLLIER, C. (2005). The impact of prudence on optimal prevention. *Economic Theory*, 26(4):989–994.
- EENPI (2010). Markets and households on low incomes. Rapport technique, Europe Economics and New Policy Institute / Office of Fair Trading.
- ELZINGA, K. G. et MILLS, D. E. (2011). The lerner index of monopoly power : origins and uses. *American Economic Review*, 101(3):558–64.
- FFRENCH-DAVIS, R. (1976). *Income Distribution in Latin America*, chapitre Policy tools and objectives of redistribution, pages 107–133. Foxley, A. and Rioseco, A.F. eds. CUP Archive.
- FOURNIER, J. et PIKETTY, T. (2015a). Generalized pareto curves : Theory and application using income and inheritance tabulations for france 1901-2012.
- GABSZEWICZ, J. J. et THISSE, J.-F. (1979). Price competition, quality and income disparities. *Journal of economic theory*, 20(3):340–359.
- GASULL, A. et UTZET, F. (2014). Approximating mills ratio. *Journal of Mathematical Analysis and Applications*, 420(2):1832–1853.
- GEISER, U., BOTTAZZI, P., EPPRECHT, M., FOKOU, G., FRITSCHI, A., RAMAKUMAR, R., SHAHBAZ, B., STEINMANN, B. et STRASSER, B. (2011). Access to livelihood assets : inclusion, exclusion, and the reality of development interventions. *Research for sustainable development : Foundations, experiences, and perspectives*, pages 313–330.

- HAMILTON, K. (2009). Low-income families : experiences and responses to consumer exclusion. *International journal of sociology and social policy*, 29(9/10):543–557.
- HE, Y. et SUN, G. (2006). Income dispersion and price discrimination. *Pacific Economic Review*, 11(1):59–74.
- IBRAGIMOV, M., IBRAGIMOV, R., KATTUMAN, P. et MA, J. (2018). Income inequality and price elasticity of market demand : the case of crossing lorenz curves. *Economic Theory*, 65(3):729–750.
- INADA, K.-I. (1963). On a Two-Sector Model of Economic Growth : Comments and a Generalization. *The Review of Economic Studies*, 30(2):119–127.
- JABLECKI, J. (2007). Continuum in economics. on the significance of the realism of assumptions in economic theory. Rapport technique, Working paper, available at : <http://www2.gcc.edu/dept/econ/ASSC/Papers2007>.
- KALECKI, M. (1945). On the gilbrat distribution. *Econometrica*, 13(2):161–170.
- KIMBALL, M. S. (1990). Precautionary saving in the small and in the large. *Econometrica*, 58(1):53–73.
- LERNER, A. P. (1934). The concept of monopoly and the measurement of monopoly power. *The Review of Economic Studies*, 1(157):75.
- LUBRANO, M. (2013). The econometrics of inequality and poverty. lecture 4 : Lorenz curves, the gini coefficient and parametric distributions. *Manuscript available online at <http://www.vcharite.univ-mrs.fr/PP/lubrano/poverty.htm>*.
- MALUEG, D. A. et SCHWARTZ, M. (1994). Parallel imports, demand dispersion, and international price discrimination. *Journal of international Economics*, 37(3-4):167–195.
- MASKIN, E. (1978). A theorem on utilitarianism. *The Review of Economic Studies*, 45(1):93–96.

- MENEZES, C., GEISS, C. et TRESSLER, J. (1980). Increasing downside risk. *The American Economic Review*, 70(5):921–932.
- NDOUR, M. et BOIDIN, B. (2012). «l'accès aux biens et services essentiels» : une notion centrale et ambiguë du développement. *L'Homme la Societe*, (3):223–248.
- OLSEN, E. O. et ROGERS, D. L. (1991). The welfare economics of equal access. *Journal of public economics*, 45(1):91–105.
- PENCHANSKY, R. et THOMAS, J. W. (1981). The concept of access : definition and relationship to consumer satisfaction. *Medical care*, 19(2):127–140.
- RIBOT, J. C. et PELUSO, N. L. (2003). A theory of access. *Rural sociology*, 68(2):153–181.
- RILEY, J. (2016). Utilitarianism and economic theory. *The New Palgrave Dictionary of Economics*, pages 1–15.
- SCHIFF, M. (2004). *On the inefficiency of inequality*. The World Bank.
- SHAFFER-LANDAU, R. (2016). *Utilitarianism*, chapitre 48, pages 417–422. John Wiley & Sons.
- SILVER, H. (1994). Social exclusion and social solidarity : three paradigms. *Int'l Lab. Rev.*, 133:531.
- SON, H. H. (2013). Inequality of human opportunities in developing asia. *Asian Development Review*, 30(2):110–130.
- TIROLE, J. (1988). *The theory of industrial organization*. MIT press.
- TRBANC, M. *et al.* (1996). Social exclusion : the concept and data indicating exclusion in slovenia. *Družboslovne razprave*, (22-23):99–114.
- VARIAN, H. R. (1984). *Microeconomic analysis, third edition*. WWNorton & Co, New-York.
- WESTON, J. F. (1950). A generalized uncertainty theory of profit. *American Economic Review*, 40(1):40–60.

WESTON, J. F. (1954). The profit concept and theory : a restatement. *Journal of Political Economy*, 62(2):152–170.

YEW-KWANG, N. (1983). *Welfare Economics : Introduction and Development of Basic Concepts*. Springer.

Appendix

3.A Annexes

3.A.1 Prix optimal qui maximise le profit

Les conditions de premier ordre par rapport à p .

$$\begin{aligned}\frac{\partial \Pi(p)}{\partial p} &= \frac{\partial}{\partial p} \left(\int_{R_0}^{+\infty} (p - c) n(R) dR \right) \\ &= N - N(R_0) - (p - c) n(R_0) \frac{\partial R_0}{\partial p}\end{aligned}\tag{3.67}$$

Le prix optimal qui maximise le profit est tel que :

$$N - N(R_0) - (p - c) n(R_0) \frac{\partial R_0}{\partial p} = 0$$

Ce qui permet d'écrire :

$$\begin{aligned}(p - c) n(R_0) \frac{\partial R_0}{\partial p} &= N - N(R_0) \\ \frac{p - c}{p} &= \frac{1}{p} \left(\frac{N - N(R_0)}{n(R_0) \frac{\partial R_0}{\partial p}} \right).\end{aligned}$$

D'après l'inégalité (3.5), $R_0 = p \left(1 + \frac{1}{u}\right)$ et donc :

$$p \left(\frac{\partial R_0}{\partial p} \right) = R_0 = p \left(1 + \frac{1}{u}\right).\tag{3.68}$$

En supposant que le prix optimal p^* est une solution intérieure, on obtient finalement

$$\frac{p^* - c}{p^*} = \frac{N - N(R_0)}{n(R_0)R_0} \equiv \frac{1}{\varepsilon_{D/p}}. \quad (3.69)$$

L'expression de l'élasticité prix de la demande agrégée $\varepsilon_{D/p}$ est telle que

$$\begin{aligned} \varepsilon_{D/p} &\equiv \frac{p}{D} \left(\frac{-dD}{dp} \right) = -\frac{p}{D} \frac{d}{dp} \left(\int_{R_0}^{+\infty} n(R) dR \right) \\ &= -\frac{p}{D} \left[-n(R_0) \left(\frac{dR_0}{dp} \right) \right] = \frac{p}{D} n(R_0) \left(\frac{dR_0}{dp} \right). \end{aligned}$$

En utilisant les équations (3.6), l'équation (3.68) et la proposition (3.8), on obtient l'expression finale de $\varepsilon_{D/p}$:

$$\begin{aligned} \varepsilon_{D/p} &= \frac{n(R_0)}{N - N(R_0)} \left(p \frac{dR_0}{dp} \right) \\ &= \frac{n(R_0)}{N - N(R_0)} R_0. \\ &= \left(\frac{pR_0'(p)}{R_0(p)} \right) \left(\frac{R_0(p)g(R_0)}{1 - G(R_0)} \right). \end{aligned} \quad (3.70)$$

3.A.2 Conditions Revenu minimum du ménage préférant la qualité unitaire X_j à la qualité unitaire X_i

Lemme. Soit $R_{i,0}^c = (1 + 1/w_i)p_i$ le revenu critique en dessous duquel un ménage préfère renoncer au bien X qu'à acheter la qualité i . Si $q_j > q_i$, dans ce cas $R_{j,0}^c > R_{i,0}^c$. A l'équilibre, on a

$$p_j > \left(\frac{1 + 1/u_i}{1 + 1/u_j} \right) p_i$$

donc $p_j > p_i$. Alors le revenu critique $R_{j,i}^c(p_j, p_i)$ au-dessus duquel la qualité q_j est préférée à la qualité q_i vérifie $R_{j,i}^c(p_j, p_i) > R_{j,0}^c$.

Démonstration. Supposons que $q_j > q_i$ et $R_{j,0}^c \leq R_{i,0}^c$ et cherchons une contradiction. L'inégalité $R_{j,0}^c \leq R_{i,0}^c$ est équivalente à

$$\left(1 + \frac{1}{u_j} \right) p_j \leq \left(1 + \frac{1}{u_i} \right) p_i,$$

où $u_j > u_i$ (de $q_j > q_i$). Ce qui peut se réécrire sous la forme

$$p_j - p_i \leq \frac{p_i}{u_i} - \frac{p_j}{u_j}.$$

Par suite,

$$u_i u_j (p_j - p_i) \leq u_j p_i - u_i p_j$$

ou

$$u_i (1 + u_j) p_j \leq u_j (1 + u_i) p_i.$$

En ajoutant aux deux membres $u_j (1 + u_j) p_j$ et en ré-arrangeant, on a

$$u_j (1 + u_j) p_j - u_j (1 + u_i) p_i \leq (u_j - u_i) (1 + u_j) p_j$$

pour obtenir finalement

$$R_{j,i}^c(p_j, p_i) = \frac{(1 + u_j) p_j - (1 + u_i) p_i}{u_j - u_i} \leq \left(1 + \frac{1}{u_j}\right) p_j = R_{j,0}^c,$$

où $R_{j,i}^c(p_j, p_i)$ dénote le revenu critique au-dessus duquel la qualité q_j est préférée à la qualité q_i .

Il en résulte que, si $R < R_{j,0}^c \leq R_{i,0}^c$, l'agent préfère renoncer au bien X ; si $R \geq R_{j,0}^c$ l'acheteur acquiert la qualité q_j et jamais la qualité q_i . Observons que tant que $R \geq R_{i,0}^c = \sup\{R_{i,0}^c, R_{j,0}^c\}$, les deux qualités sont considérées comme importantes à acquérir. Dans le cas où $R \geq R_{j,i}^c(p_j, p_i)$, la qualité q_j apparaît plus intéressante que la qualité q_i .

Clairement, il peut être le cas où la qualité q_i n'est jamais acquise à l'équilibre. Par conséquent, il nous faut avoir $R_{j,0}^c > R_{i,0}^c$. A partir de l'expression de $R_{k,0}^c = (1 + 1/u_k) p_k$, $k = i, j$, nous obtenons

$$p_j > \left(\frac{1 + 1/u_i}{1 + 1/u_j}\right) p_i$$

et puisque $u_i < u_j$, il s'ensuit que $p_j > p_i$. La démarche suivie précédemment permet que

conclure que l'inégalité $R_{j,0}^c > R_{i,0}^c$ implique que

$$R_{j,i}^c(p_j, p_i) > R_{j,0}^c.$$

Avec la disponibilité de deux qualités, le bien X n'est jamais acquis quand $R \in [0, R_{1,0}^c[$, la qualité q_1 est acquise quand $R \in [R_{1,0}^c, R_{2,1}^c[$, et la qualité q_2 tant que $R \geq R_{2,1}^c$. \square

3.A.3 Prix optimal qui maximise le bien-être social

On considère l'équation 3.9 définissant le bien-être social. La *CPO* en p est telle que

$$\begin{aligned} \frac{\partial W(p)}{\partial p} &= \frac{\partial}{\partial p} \left(\left\{ \int_0^{R_0} \ln(R) n(R) dR + \int_{R_0}^{+\infty} [\ln(1+u) + \ln(R-p)] n(R) dR \right\} \right. \\ &\quad \left. + \left\{ \int_{R_0}^{+\infty} (p-c) n(R) dR \right\} \right) \\ &= [\ln(R_0) - \ln(1+u) - \ln(R_0-p) - (p-c)] n(R_0) \frac{\partial R_0}{\partial p} \\ &\quad + \int_{R_0}^{+\infty} \left[1 - \frac{1}{R-p} \right] n(R) dR. \end{aligned}$$

On sait que quand $R = R_0$, un ménage est exactement indifférent entre acheter le bien ou ne pas l'acheter. Donc

$$U(x=1, R_0) = \ln(1+u) + \ln(R_0-p) = \ln(R_0) = U(x=0, R_0).$$

Par suite :

$$\frac{\partial W(p)}{\partial p} = -(p-c) n(R_0) \frac{\partial R_0}{\partial p} + \int_{R_0}^{+\infty} \left[1 - \frac{1}{R-p} \right] n(R) dR. \quad (3.71)$$

On peut vérifier que le prix optimal p qui maximise le profit ne maximise pas le bien-être social. A partir de la *CPO* de l'équation (3.10) de profit, on déduit : $N - N(R_0) =$

$(p - c) n(R_0) \frac{\partial R_0}{\partial p}$ et donc

$$\begin{aligned}
\frac{\partial W(p)}{\partial p} &= -[N - N(R_0)] + \int_{R_0}^{+\infty} \left[1 - \frac{1}{R-p}\right] n(R) dR \\
&= -[N - N(R_0)] + [N - N(R_0)] - \int_{R_0}^{+\infty} \left[\frac{1}{R-p}\right] n(R) dR \\
&= - \int_{R_0}^{+\infty} \left[\frac{1}{R-p}\right] n(R) dR < 0.
\end{aligned} \tag{3.72}$$

Quelques manipulations permettent de déterminer le prix socialement optimal p^* définissant la marge qui maximise le bien-être social :

$$\begin{aligned}
(p^* - c) n(R_0^*) \frac{\partial R_0^*}{\partial p} &= \int_{R_0^*}^{+\infty} \left[1 - \frac{1}{R-p^*}\right] n(R) dR \\
\frac{p^* - c}{p^*} &= \frac{1}{p^*} \frac{\int_{R_0^*}^{+\infty} \left[1 - \frac{1}{R-p^*}\right] n(R) dR}{n(R_0^*) \frac{\partial R_0^*}{\partial p}}.
\end{aligned} \tag{3.73}$$

L'expression de $\varepsilon_{D/p}$ obtenue en 3.12 permet de simplifier l'équation (3.73) :

$$\begin{aligned}
\frac{p^* - c}{p^*} &= \frac{\int_{R_0^*}^{+\infty} \left[1 - \frac{1}{R-p^*}\right] n(R) dR}{p^* n(R_0^*) \frac{\partial R_0^*}{\partial p}} \\
&= \frac{\int_{R_0^*}^{+\infty} \left[1 - \frac{1}{R-p^*}\right] \frac{n(R)}{N - N(R_0^*)} dR}{\frac{n(R_0^*)}{N - N(R_0^*)} \left(p^* \frac{\partial R_0^*}{\partial p}\right)} \\
&= \frac{1 - \int_{R_0^*}^{+\infty} \left(\frac{1}{R-p^*}\right) \frac{n(R)}{N - N(R_0^*)} dR}{\varepsilon_{D/p}}.
\end{aligned} \tag{3.74}$$

3.A.4 Caractérisation d'une fonction $R_0(p)$, d'une fonction $D(p)$ et établissement d'une condition de non-exclusion

Proposition 3.7. Pour une fonction d'utilité indirecte telle que définie à l'équation (3.62) associée à $\mathcal{U}(X, R)$ et satisfaisant les hypothèses (3.1) et (3.3), il existe un revenu critique $R_0(p, u) > p$, tel que tous les ménages ayant un revenu supérieur (inférieur) à ce revenu critique auront (n'auront pas) accès au bien X . De plus, ce revenu seuil augmente

plus que proportionnellement avec p et décroît avec u . Formellement :

$$\frac{dR_0}{dp} > 1 \quad \text{and} \quad \frac{dR_0}{du} < 0.$$

Démonstration. Proposition 3.7 : Annexe 3.A.4 □

Soit $b(R)$ le bénéfice net retiré de l'acquisition du bien X et donc :

$$b(R) = [u + v(R - p)] - v(R).$$

Par conséquent,

$$\frac{d}{dR} [b(R)] = v'(R - p) - v'(R) \geq 0$$

étant donné que $v'(\cdot)$ est décroissante ($v(\cdot)$ est concave). De plus,

$$\lim_{R \rightarrow p^+} \{[u + v(R - p)] - v(R)\} = -\infty.$$

En d'autres termes, un ménage n'achète jamais le bien quand son revenu se rapproche du prix p du bien. Par contre, $\forall p, u > 0$, quand le revenu est suffisant, on préfère toujours acheter le bien. En effet, comme la fonction $v(R)$ est concave, on en déduit que (figure 3.5.4) :

$$pv'(R_0 - p) \leq [u + v(R_0 - p)] - v(R_0) = u - [v(R_0) - v(R_0 - p)] \leq u - pv'(R_0)$$

soit

$$pv'(R_0 - p) \leq [u + v(R_0 - p)] - v(R_0) \leq u - pv'(R_0).$$

En passant à la limite, on obtient :

$$u \leq \lim_{R \rightarrow +\infty} \{[u + v(R_0 - p)] - v(R_0)\} \leq u.$$

Ce qui prouve qu'il *existe un revenu minimal* $R_0(p) > p$ tel que

$$[u + v(R_0 - p)] = v(R_0) \quad (3.75)$$

vérifiant

$$v(R - p) \leq v(R) \quad \forall R / 0.8 \leq R \leq R_0,$$

$$v(R - p) \geq v(R) \quad \forall R \geq R_0.$$

Ainsi, tout ménage qui a ce revenu R_0 est indifférent entre l'achat ou non du bien. On sait de plus que si $R > R_0$, on achète toujours le bien - et on ne l'achète pas dans le cas contraire.

Par ailleurs, comme R est un bien essentiel, on peut aisément vérifier que son augmentation procure une utilité de moins en moins importante retirée de la consommation d'une unité du bien X (hypothèse de décroissance de l'utilité marginale du revenu). Ceci se vérifie en différenciant l'équation (3.75) par rapport à u :

$$1 + \frac{dR_0}{du} v'(R_0 - p) = \frac{dR_0}{du} v'(R_0)$$

de sorte que

$$\frac{dR_0}{du} = -[v'(R_0 - p) - v'(R_0)]^{-1} < 0.$$

Par contre, quand le prix p s'accroît, le revenu R_0 donnant accès au bien aux ménages intéressés, R_0 étant une fonction de p , augmente aussi. L'augmentation induite de R_0 est plus que proportionnelle à l'accroissement de p du fait que le ménage valorise davantage le bien en ce sens qu'il retire une utilité plus importante, proportionnelle à l'augmentation du revenu, à l'achat d'une unité du bien. Cela se démontre en différenciant l'équation (3.75) cette fois-ci par rapport à p :

$$\left(\frac{dR_0}{dp} - 1 \right) v'(R_0 - p) = \frac{dR_0}{dp} v'(R_0)$$

ce qui permet d'écrire

$$\frac{dR_0}{dp} = \frac{v'(R_0 - p)}{v'(R_0 - p) - v'(R_0)} > 1. \quad (3.76)$$

Proposition 3.8 On suppose que $R_0 \in [\underline{R}; \bar{R}]$. Le revenu $R_0(p)$ est une fonction croissante de p et la demande $D(p)$ est décroissante en p et définie par :

$$D(p) \equiv \int_{R_0}^{+\infty} n(R) dR = N - N(R_0), \text{ pour tout } p \geq \underline{p}.$$

$$= \begin{cases} N(1-m)[1-G(R_0)] & \text{pour tout } p \geq \underline{p}. \\ N & \text{sinon.} \end{cases}$$

où $n(R) : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ est la fonction de densité du nombre d'individus ayant un revenu R et $n(R) = N(1-m)g(R)$, $g(R)$ étant la fonction de densité associée à $G(R)$.

Démonstration. On a établi en (3.7) que R_0 est une fonction croissante de p . Clairement, tant que $R_0 \in [\underline{R}; \bar{R}]$, on a :

$$\frac{\partial D}{\partial p} = -n(R_0) \frac{dR_0}{dp} < 0.$$

Par suite, l'élasticité-prix de la demande est donnée par :

$$\begin{aligned} \varepsilon_{D/p} &= \left(\frac{p}{D}\right) \left(-\frac{\partial D}{\partial p}\right) \\ &= \frac{p}{\int_{R_0}^{+\infty} n(R) dR} n(R_0) \frac{dR_0}{dp} \\ &= p \frac{n(R_0)}{\int_{R_0}^{+\infty} n(R) dR} \left[\frac{v'(R_0 - p)}{v'(R_0 - p) - v'(R_0)} \right]. \\ &= p \frac{n(R_0)}{\int_{R_0}^{+\infty} n(R) dR} \left[\frac{v'(R_0 - p)}{v'(R_0 - p) - v'(R_0)} \right]. \end{aligned} \quad (3.77)$$

Ou de manière équivalente

$$\begin{aligned}
\varepsilon_{D/p} &= \left(\frac{p}{D} \right) \left(- \frac{\partial D}{\partial p} \right) \\
&= - \frac{p}{N(1-m)[1-G(R_0(p))]} \frac{\partial D}{\partial p} [N(1-m)[1-G(R_0(p))]] \\
&= \frac{pR'_0(p)G'(R_0(p))N(1-m)}{N(1-m)[1-G(R_0(p))]} \\
&= \left(\frac{pR'_0(p)}{R_0(p)} \right) \left(\frac{R_0(p)g(R_0)}{1-G(R_0)} \right), \quad G'(R_0(p)) \equiv g(R_0)
\end{aligned}$$

□

Proposition 3.9. On note c le coût unitaire de fabrication du bien X . Les profits du monopole s'écrivent donc :

$$\Pi(p) = (p - c) D(p, R).$$

Soit $F(p, R)$ une fonction continue et dérivable sur $\mathbb{R} \times \mathbb{R}_+^n$ définie par :

$$F(p, R) \equiv \frac{\partial \Pi(p)}{\partial p} = D(p, R) + (p - c) \frac{\partial D(p, R)}{\partial p}$$

et soit $p^*(\mathbf{R}) : \mathbb{R}_+^n \rightarrow \mathbb{R}$ la fonction implicite définie par $F(p, R) \equiv 0$. Si $F(p, \cdot)$ est Schur-convexe (resp. concave) pour tout $p \in \mathbb{R}$ et $[\partial F(p, R) / \partial p] < 0$ pour tout $\mathbf{R} \in \mathbb{R}_+^n$ alors la fonction $p^*(\mathbf{R})$ est Schur-concave (resp. convexe).

Démonstration. Soit $(\mathbf{R}, \mathbf{R}') \in (\mathbb{R}_+^n)^2$ tel que $\mathbf{R} \prec \mathbf{R}'$. Soit $(p, p') \in \mathcal{P}^2$ tel que $F(p, \mathbf{R}) = 0$ et $F(p', \mathbf{R}') = 0$. Comme $F(p, \cdot)$ est Schur-convexe (resp. concave), alors

$$0 = F(p, \mathbf{R}) \leq (\geq) F(p, \mathbf{R}'), \quad \forall p \implies$$

$$0 = F(p', \mathbf{R}) \leq (\geq) F(p', \mathbf{R}'), \quad \forall p'.$$

et donc

$$F(p', \mathbf{R}') = 0 = F(p, \mathbf{R}) \leq (\geq) F(p, \mathbf{R}') \tag{3.78}$$

ou encore

$$F(p', \mathbf{R}) \leq (\geq) F(p', \mathbf{R}') = 0 = F(p, \mathbf{R}). \tag{3.79}$$

Puisque $[\partial F(p, R) / \partial p] < 0$, les inégalités (3.78) et (3.79) impliquent que

$$p' \leq (\geq) p.$$

Par conséquent, la fonction implicite $p^*(\mathbf{R})$ définie par $F(p, R) \equiv 0$ est Schur-concave (resp. convexe). \square

Proposition 3.10. On note c le coût unitaire de fabrication du bien X . Les profits du monopole s'écrivent donc :

$$\Pi(p) = (p - c) [N - N(R_0)]. \quad (3.80)$$

Si $R_0 \leq \underline{R}$ (il n'y a pas d'exclusion), le monopole qui maximise son profit a intérêt à adopter un prix p tel que $R_0 = \underline{R}$.

Démonstration. S'il n'y a pas d'exclusion ($\underline{R} \geq R_0$), le profit s'écrit :

$$\Pi(p) = (p - c) [N - N(\underline{R})]. \quad (3.81)$$

Donc la CPO de la maximisation du profit s'écrit

$$\frac{\partial \Pi}{\partial p} = N - N(\underline{R}) > 0.$$

Donc, tant qu'il n'y a pas exclusion, le monopole a intérêt à augmenter le prix p de sorte que R_0 va augmenter jusqu'à atteindre \underline{R} . \square

Proposition 3.11. Le prix unique p vérifiant $R_0(p) = \underline{R}$ est déterminé dans le cadre d'une distribution des revenus telle que le monopole trouve profitable de ne pas exclure de consommateurs.

Démonstration. Cette preuve s'établit en trois étapes :

a) Condition de non-exclusion. Pour que le monopole ne trouve pas profitable de faire de l'exclusion, il faut qu'il aie envie de baisser les prix p lorsque $R_0 \geq \underline{R}$. Donc en

particulier quand, en référence à l'équation (3.81),

$$\left. \frac{\partial \Pi}{\partial p} = N - N(\underline{R}) - (p - c) n(\underline{R}) \frac{dR_0}{dp} \right]_{R_0=\underline{R}} \leq 0.$$

Cette condition se ré-écrit :

$$\frac{p - c}{p} \geq \frac{1}{\varepsilon_{D/p}(\underline{R})}$$

où $\varepsilon_{D/p}(\underline{R}) = \left. \frac{\partial D_0}{\partial R_0} \right]_{R_0=\underline{R}} \times \frac{R_0}{D_0}$. On déduit de cette dernière inégalité la condition de non-exclusion :

$$(p - c) n(\underline{R}) \left. \frac{dR_0}{dp} \right]_{R_0=\underline{R}} \geq N - N(\underline{R}). \quad (3.82)$$

b) Le prix p qui maximise le profit est tel que $R_0(p) = \underline{R}$. On cherche à déterminer et à évaluer l'impact d'une variation du revenu minimum \underline{R} sur le prix p , $dp/d\underline{R}$. Comme p est tel que $R_0(p) = \underline{R}$ donc

$$\begin{aligned} \frac{dR_0(p)}{d\underline{R}} &= \frac{d\underline{R}}{d\underline{R}} \\ \implies \left. \frac{dR_0}{dp} \right]_{R_0=\underline{R}} \frac{dp}{d\underline{R}} &= 1 \end{aligned}$$

et donc, en utilisant l'équation (??),

$$\begin{aligned} \frac{dp}{d\underline{R}} &= \left[\left. \frac{dR_0}{dp} \right]_{R_0=\underline{R}} \right]^{-1} \\ &= \left(\frac{v'(R_0 - p) - v'(R_0)}{v'(R_0 - p)} \right) \Bigg|_{R_0=\underline{R}} < 1. \end{aligned}$$

c) Enfin, on définit une condition générale garantissant un accès au bien à tous. Lorsque la maximisation du profit donne lieu à exclusion ($R_0 > \underline{R}$) alors on a une solution intérieure au problème de maximisation du profit définie par la CPO obtenue à l'équation (3.10) :

$$\frac{\partial \Pi}{\partial p} = [N - N(R_0)] - (p - c) n(R_0) \frac{\partial R_0}{\partial p}.$$

De cette *CPO*, la condition pour que l'entreprise veuille baisser son prix est telle que :

$$(p - c) n(R_0) \left. \frac{dR_0}{dp} \right|_{R_0} \geq N - N(R_0). \quad (3.83)$$

Cette condition se ré-écrit :

$$\begin{aligned} N - N(R_0) &\leq (p - c) n(R_0) \left. \frac{dR_0}{dp} \right|_{R_0} \\ \frac{N - N(R_0)}{n(R_0)} &\leq (p - c) \left. \frac{dR_0}{dp} \right|_{R_0} \\ \frac{n(R_0)}{N - N(R_0)} &\geq \left[(p - c) \left. \frac{dR_0}{dp} \right|_{R_0} \right]^{-1}. \end{aligned}$$

En utilisant le résultat de la relation établie en (3.75), cette inégalité se ré-écrit :

$$h(R_0) \equiv \frac{n(R_0)}{N - N(R_0)} \geq \frac{v'(R_0 - p) - v'(R_0)}{(p - c)v'(R_0 - p)} \equiv w(R_0) \quad (3.84)$$

où R_0 est tel que $[u + v(R_0 - p)] = v(R_0)$ (équation 3.75). Cette condition lie donc la distribution des revenus à travers la fonction du hasard $h(R_0) = \{n(R_0) / [N - N(R_0)]\}$ aux préférences des ménages à travers l'expression

$$w(R_0) = [v'(R_0 - p) - v'(R_0)] / [(p - c)v'(R_0 - p)].$$

□

Proposition 3.12. Sous l'hypothèse additionnelle que $v'''(.) > 0$ ²³, la fonction

$$w(R_0) = \left(\frac{1}{p-c} \right) \frac{v'(R_0-p) - v'(R_0)}{v'(R_0-p)}, \quad R_0(p) > p > c$$

est monotone décroissante.

Démonstration. Il s'agit de démontrer que $w'(R_0) < 0, \forall R_0(p) > p > c$.

On rappelle que $R_0 = r_0(p)$ ce qui implique que $p = g(R_0)$, $g = r_0^{-1}$. Et donc,

$$\begin{aligned} \frac{dp}{dR_0} &\equiv g'(R_0) \\ &= \frac{1}{r_0'(g(R_0))} \\ &= \frac{1}{r_0'(p)} \end{aligned}$$

ce que l'on peut noter sans risque de confusion, par

$$\frac{dp}{dR_0} = \frac{1}{R_0'(p)}.$$

Soit alors

$$\begin{aligned} m(R_0) &= \frac{1}{p-c} = \frac{1}{g(R_0)-c}, \\ n(R_0) &= \frac{v'(R_0-p) - v'(R_0)}{v'(R_0-p)}. \end{aligned}$$

23. Dans la littérature traitant de l'analyse du risque, la dérivée d'ordre trois est utilisée pour caractériser le comportement des individus dits « averses au risque de pertes potentielles » (downside risk aversion). Selon un article publié par Menezes *et al.* (1980), ce type d'aversion au risque décrit le comportement attribué à un motif d'évitement d'une catastrophe. Ces individus sont alors ceux qui évitent des situations risquées offrant de potentiels gains substantiels du fait que ces situations rendent ces individus vulnérables à des pertes en dessous d'un certain seuil. Eeckhoudt et Gollier (2005); Kimball (1990) qualifient ces individus averses au risque de « prudents » c'est-à-dire des individus qui, s'étant obligés à faire face à un risque accru, préfèrent lier ce risque aux meilleurs résultats qu'aux pires. La caractérisation du comportement « prudent » est établi à travers le corollaire 1 stipulant qu'un individu est prudent si et seulement si son utilité espérée vérifie $u'''(.) > 0$. Différentes recherches tant théoriques qu'empiriques utilisent ce résultat caractérisant le comportement des individus prudents (Deck et Schlesinger, 2018; Crainich *et al.*, 2015).

Les dérivées de ces deux fonctions sont telles que

$$m'(R_0) = -\frac{1}{R'_0(p)(p-c)^2}$$

et

$$\begin{aligned} n'(R_0) &= \frac{v'(R_0-p) \left\{ v''(R_0-p) \left[1 - \frac{1}{R'_0(p)} \right] - v''(R_0) \right\}}{[v'(R_0-p)]^2} \\ &\quad - \frac{v''(R_0-p) \left[1 - \frac{1}{R'_0(p)} \right] [v'(R_0-p) - v'(R_0)]}{[v'(R_0-p)]^2} \\ &= \frac{v'(R_0-p) v''(R_0-p) [R'_0(p) - 1] - v'(R_0-p) v''(R_0) R'_0(p)}{R'_0(p) [v'(R_0-p)]^2} \\ &\quad - \frac{v'(R_0-p) v''(R_0-p) [R'_0(p) - 1] - v'(R_0) v''(R_0-p) [R'_0(p) - 1]}{R'_0(p) [v'(R_0-p)]^2} \\ &= \frac{v'(R_0) v''(R_0-p) [R'_0(p) - 1] - v'(R_0-p) v''(R_0) R'_0(p)}{R'_0(p) [v'(R_0-p)]^2}. \end{aligned}$$

La dérivée de la fonction $w(R_0)$ s'écrit alors :

$$\begin{aligned} w'(R_0) &= - \left[\frac{1}{R'_0(p)(p-c)^2} \right] \left[\frac{v'(R_0-p) - v'(R_0)}{v'(R_0-p)} \right] \\ &\quad + \left(\frac{1}{p-c} \right) \left[\frac{v'(R_0) v''(R_0-p) [R'_0(p) - 1] - v'(R_0-p) v''(R_0) R'_0(p)}{R'_0(p) [v'(R_0-p)]^2} \right] \\ &= w'_1(R_0) + w'_2(R_0). \end{aligned}$$

On peut voir que $w'_1(R_0) < 0$ car $R_0(p) > p > c$ et

$$0 < \frac{v'(R_0-p) - v'(R_0)}{v'(R_0-p)} < 1.$$

Alors $w'(R_0) < 0$ que si $w'_2(R_0) < 0$ c'est-à-dire que si

$$v'(R_0) v''(R_0-p) [R'_0(p) - 1] - v'(R_0-p) v''(R_0) R'_0(p) < 0$$

ou de manière équivalente que si

$$\frac{v''(R_0 - p)}{v''(R_0)} \left[\frac{R'_0(p) - 1}{R'_0(p)} \right] < \frac{v'(R_0 - p)}{v'(R_0)}. \quad (3.85)$$

Comme

$$0 < \frac{R'_0(p) - 1}{R'_0(p)} < 1,$$

une condition suffisante pour que l'inégalité (3.85) soit toujours vraie est que

$$\frac{v''(R_0 - p)}{v''(R_0)} \leq \frac{v'(R_0 - p)}{v'(R_0)}.$$

Par hypothèse $v''(\cdot) < 0$ alors $v'(R_0 - p) > v'(R_0)$ et donc

$$\frac{v'(R_0 - p)}{v'(R_0)} > 1. \quad (3.86)$$

Par ailleurs, l'hypothèse $v'''(\cdot) > 0$ implique que $v''(R_0) > v''(R_0 - p)$ ou de manière équivalente

$$\frac{v''(R_0 - p)}{v''(R_0)} < 1. \quad (3.87)$$

En conséquence, par les inégalités (3.86) et (3.87),

$$\frac{v'(R_0 - p)}{v'(R_0)} > 1 > \frac{v''(R_0 - p)}{v''(R_0)}.$$

Par suite

$$\frac{v''(R_0 - p)}{v''(R_0)} < \frac{v'(R_0 - p)}{v'(R_0)}.$$

La fonction $w(R_0) = [1/(p - c)][v'(R_0 - p) - v'(R_0)]/v'(R_0 - p)$ est donc décroissante.

□

Les fonctions $h(R_0)$ et $w(R_0)$ sont donc deux fonctions décroissantes²⁴. On s'attend à ce qu'à mesure que R_0 augmente, $h(R_0)$ et $w(R_0)$ diminuent progressivement jusqu'à ce que, pour tout R_0 , $h(R_0) \geq w(R_0)$ traduisant ainsi l'élimination de l'exclusion.

24. La fonction $h(R)$ peut être décroissante, constante ou croissante. Elle vérifie les propriétés suivantes : $h(R) \geq 0$ et $\int_0^{+\infty} h(R) dR \rightarrow +\infty$.

Conclusion Générale

Cette thèse, constituée de trois essais, tente d'apporter quelques maigres éléments de réponse à la problématique de la pauvreté caractérisée par sa multidimensionnalité.

Dans le cadre d'un premier essai, nous avons abordé de manière très partielle, il faut l'admettre, la problématique de l'inférence statistique, en utilisant des méthodes statistiques dites de "capture-recapture", relative aux populations difficiles à enquêter en particulier les ménages vivant en extrême pauvreté. La lutte contre la pauvreté passe d'abord par l'identification et le dénombrement des pauvres. Nous avons donc développé une extension de ces méthodes statistiques nous permettant de re-évaluer l'estimateur de Lincoln-Peterson et sa variance en le comparant à un estimateur naïf pour la taille de la population. Cette extension est utilisée pour estimer deux probabilités conditionnelles qui ont servi à l'estimation du modèle de régression binaire $P(y_i = 1) = \beta_0 + \beta_1 x_i + \varepsilon_i$ où $x \in \{0, 1\}$ et $y \in \{0, 1\}$ lorsque le biais des estimateurs naïfs ne peut être considéré comme étant nul et que la taille des échantillons statistiques est suffisante pour que la prise en compte de l'hétérogénéité puisse être considérée comme bénéfique. Ce chapitre de thèse ouvre une perspective d'élaboration d'outils d'inférence pour les populations difficiles à joindre en général. En perspective, nous tâcherons de généraliser les résultats d'estimation de cette régression binaire dans un cadre de la disponibilité de trois listes et pour un nombre K quelconque de variables indépendantes discrètes et continues. Nous envisageons de procéder, dans ce cadre, au dénombrement des enfants de rue en Haïti et d'en analyser leurs principales caractéristiques.

A travers le deuxième essai, la problématique de l'arrosage à petite échelle a été traitée en développant un modèle intégrant deux risques climatiques: la sécheresse et l'inondation. Ce modèle permettant d'évaluer comment cette activité d'arrosage contribue à améliorer les revenus des petits exploitants agricoles touchés en général par la pauvreté. Cet essai tente ainsi d'apporter une contribution à la problématique de lutte contre la pauvreté rurale d'importance dans les pays en développement. Les résultats de ce modèle peuvent être améliorés en modélisant l'installation, de manière graduelle et non pas de manière dichotomique, d'une sécheresse et/ou d'une inondation. En perspective, on concevra un cadre incitatif à l'arrosage et on analysera le lien entre arrosage et impact des inondations

pour étudier les conséquences sur les incitations.

L'accès limité ou le non accès des pauvres à un certain nombre de biens et services constitue l'un des problèmes majeurs auxquels font face ces pauvres. Cette problématique a été abordée à travers une extension du modèle classique de demande unitaire en y incorporant l'effet revenu (essai numéro 3). Cette extension nous a servi à identifier comment les caractéristiques de la distribution des revenus dans une société constituent un des facteurs clé qui conditionne le niveau d'accès d'un consommateur à un bien ou un service d'une part. D'autre part, nous avons pu aussi identifier, dans le cadre de cette distribution des revenus, les conditions générales dans lesquelles un monopole trouve profitable de ne pas exclure des consommateurs. En perspective, nous chercherons, dans un premier temps, à analyser l'impact de l'institution d'un impôt redistributif sur la distribution des revenus dans l'économie. Dans un second temps, nous caractériserons les types de bien (les préférences) pour lesquels les inégalités entraînent une hausse des prix.